



**AFRL-RY-WP-TR-2012-0202**

**MULTI-YIELD RADIO FREQUENCY  
COUNTERMEASURES  
INVESTIGATIONS AND DEVELOPMENT (MYRIAD)  
Task Order 0006: Integrated Multi-Modal RF Sensing**

**Mark L. Brockman  
Dynetics, Inc.**

**Steven Kay and Quan Ding  
University of Rhode Island**

**Sean M. O'Rourke and A. Lee Swindlehurst  
University of California, Irvine**

**Antonia Papandreou-Suppappola  
Arizona State University**

**Ram M. Narayanan and Surendra S. Bhat  
The Pennsylvania State University**

**AUGUST 2012  
Final Report**

**Approved for public release; distribution unlimited**

**AIR FORCE RESEARCH LABORATORY  
SENSORS DIRECTORATE  
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7320  
AIR FORCE MATERIEL COMMAND  
UNITED STATES AIR FORCE**

## NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the USAF 88<sup>th</sup> Air Base Wing (88 ABW) Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (<http://www.dtic.mil>).

AFRL-RY-WP-TR-2012-0202 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNED//

---

CHRISTOPHER COOPER, Program Manager  
Radio Frequency Exploitation Branch

//SIGNED//

---

JEFF SANDERS, Branch Chief  
Radio Frequency Exploitation Branch

//SIGNED//

---

DOUG HAGER, Division Chief  
Layered Sensing Exploitation Division

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

\*Disseminated copies will show “//signature//” stamped or typed above the signature blocks.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. <b>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</b></p>					
1. REPORT DATE (DD-MM-YY) August 2012		2. REPORT TYPE Final		3. DATES COVERED (From - To) 18 March 2009 – 8 August 2012	
4. TITLE AND SUBTITLE MULTI-YIELD RADIO FREQUENCY COUNTERMEASURES INVESTIGATIONS AND DEVELOPMENT (MYRIAD) Task Order 0006: Integrated Multi-Modal RF Sensing				5a. CONTRACT NUMBER FA8650-08-D-1303-0006	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER 61102F	
6. AUTHOR(S) Mark L. Brockman (Dynetics, Inc.) Steven Kay and Quan Ding (University of Rhode Island) Sean M. O'Rourke and A. Lee Swindlehurst (University of California, Irvine) Antonia Papandreou-Suppappola (Arizona State University) Ram M. Narayanan and Surendra S. Bhat (The Pennsylvania State University)				5d. PROJECT NUMBER 2305	
				5e. TASK NUMBER YR	
				5f. WORK UNIT NUMBER Y0K2	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Radio Frequency Exploitation Branch Layered Sensing Exploitation Division Air Force Research Laboratory, Sensors Directorate Wright-Patterson Air Force Base, OH 45433-7320 Air Force Materiel Command, United States Air Force				8. PERFORMING ORGANIZATION REPORT NUMBER AFRL-RY-WP-TR-2012-0202	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory Sensors Directorate Wright-Patterson Air Force Base, OH 45433-7320 Air Force Materiel Command United States Air Force				10. SPONSORING/MONITORING AGENCY ACRONYM(S) AFRL/RYP	
				11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S) AFRL-RY-WP-TR-2012-0202	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.					
13. SUPPLEMENTARY NOTES This work was funded in whole or in part by Department of the Air Force contract FA8650-08-D-1303-0006. The U.S. Government has for itself and others acting on its behalf an unlimited, paid-up, nonexclusive, irrevocable worldwide license to use, modify, reproduce, release, perform, display, or disclose the work by or on behalf of the U.S. Government.. PAO Case Number 88ABW-2012-5074, Clearance Date 18 September 2012. Report contains color.					
14. ABSTRACT The Integrated Multi-Modal RF Sensing (IMMRF) effort performed basic research to examine key parameters of integrated multi-modal RF sensor design and algorithm development. A statistical framework for the program was developed, which ensured the overall statistical validity of the approaches. Dynamic waveform design for agile RF sensing, enhanced detection capabilities, and optimized tracking performance was investigated. Approaches were developed for using multiple, spatially distributed, adaptive multi-modal sensors for multiple target tracking and data association. Finally, a radar test bed was developed to support testing and refinement of theories and algorithms.					
15. SUBJECT TERMS multi-modal RF sensing, MIMO radar, statistical radar framework, agile waveform					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT: SAR	18. NUMBER OF PAGES 172	19a. NAME OF RESPONSIBLE PERSON (Monitor) Phillip Woodhull 19b. TELEPHONE NUMBER (Include Area Code) N/A
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			

# Table of Contents

<u>Section</u>	<u>Page</u>
List of Figures.....	iv
List of Tables .....	vi
1. Summary.....	1
2. Introduction.....	3
3. Statistical Framework .....	5
3.1. Methods, Assumptions, and Procedures .....	5
3.2. Exponentially Embedded Families for Multimodal Sensor Processing.....	6
3.3. Sensor Integration by Joint PDF Construction Using the EEF .....	14
3.4. Robust Signal Detection Using the EEF.....	38
3.5. Conclusions .....	47
4. Algorithm Development and Implementation.....	48
4.1. Methods, Assumptions, and Procedures .....	48
4.2. Target Dynamic Models.....	49
4.3. Sensor Measurement Models .....	50
4.4. Ideal Sensing Environment: The Extended Kalman Filter.....	53
4.5. Cluttered Sensing Environment: The MS-JPDAF.....	53
4.6. Finite Field-of-View Modeling .....	54
4.7. Proposed Closed-Loop Control & Optimization Algorithm .....	56
4.8. Simulations and Results.....	60
4.9. Conclusions .....	79
5. Tracking and Waveform Agility.....	80
5.1. Methods, Assumptions, and Procedures .....	80
5.2. Multi-Modal Track-Before-Detect for Low Observable Targets .....	82
5.3. Waveform-Agile Multi-Modal TBD .....	86
5.4. Statistical Dependent RF-EO Measurements .....	87
5.5. Multiple Target Tracking.....	88
5.6. Simulation and Experimental Results.....	96
5.7. Conclusions .....	111
6. Radar Test-Bed Development.....	112
6.1. Methods, Assumptions, and Procedures .....	112
6.2. Concept of Multimodal Radar .....	114
6.3. Results from Theory and Simulations.....	120
6.4. Field Measurement Results .....	140
6.5. Conclusions .....	152
7. Conclusions.....	154
8. References.....	156

## List of Figures

<b>Figure</b>	<b>Page</b>
Figure 1. ROC Curves for Different Detectors. ....	13
Figure 2. Distributed Detection/Classification System with Two Sensors .....	16
Figure 3. GLRT ROC Curves for Uncorrelated Gaussian Mixture Noise .....	35
Figure 4. GLRT ROC Curves for Correlated Gaussian Mixture Noise .....	35
Figure 5. Probability of Correct Classification for Both Methods .....	37
Figure 6. Probability of Correct Classification for Both Methods .....	38
Figure 7. PDF Parameters for Different Models .....	46
Figure 8. ROC curves for different detectors when $\mathbf{x} = \mathbf{A} + \mathbf{w}$ .....	46
Figure 9. ROC Curves for Different Detectors when $\mathbf{x} = \mathbf{s} + \mathbf{w}$ .....	47
Figure 10. Three-Target Ideal Sensing Scenario Tracks .....	62
Figure 11. RMS Position Error for Three-Target Ideal Scenario .....	62
Figure 12. Parameter Evolution for Three-Target Ideal Scenario .....	63
Figure 13. Target Tracks for Objective Function Comparison Analysis .....	64
Figure 14. RMS Position Errors for Objective Function Comparison .....	65
Figure 15. Parameter Evolution for Adaptive Method with Approximation Objective .....	66
Figure 16. Parameter Evolution for Adaptive Method with MII Objective .....	66
Figure 17. Target Tracks for Non-adaptive Method in Clutter .....	68
Figure 18. Target Tracks for Adaptive Method in Clutter .....	68
Figure 19. RMS Position Error for Clutter Scenario .....	69
Figure 20. RMS Velocity Error for Clutter Scenario .....	69
Figure 21. Adaptive PRF Evolution for RF Mode in Clutter .....	70
Figure 22. Adaptive Bandwidth Evolution for RF Mode in Clutter .....	70
Figure 23. Adaptive DAS Evolution for EO Mode in Clutter .....	71
Figure 24. Maneuvering Target Scenario Trajectories for Clutter Analysis .....	73
Figure 25. Probability of Non-Empty Association Comparisons .....	74
Figure 26. RMS Position Error Comparison for Target 1 .....	76
Figure 27. RMS Position Error Comparison for Target 2 .....	76
Figure 28. RMS Velocity Error Comparison for Target 1 .....	77
Figure 29. RMS Velocity Error Comparison for Target 2 .....	77
Figure 30. Probability of Non-Empty Association Comparison .....	78
Figure 31. Ambiguity Function Measurement at 6 dB SNR .....	84
Figure 32. Ambiguity Function Measurement at 40 dB SNR .....	85
Figure 33. MSE Comparison of PF-TBD Estimation Position Error, Weight = (1,1,0,0) .....	97
Figure 34. MSE Comparison of PF-TBD Estimation Position Error, Weight = (1,1,1,1) .....	98
Figure 35. MSE Comparison of PF-TBD Estimation Position Error, Weight = (0,0,1,1) .....	98
Figure 36. True and Estimated Trajectory Using PF-TBD .....	99
Figure 37. Averaged Position Error Using PF-TBD .....	99
Figure 38. RF-EO Experimental Sensing Platform at Pennsylvania State University .....	100
Figure 39. Experiment Field View: Left View of the Platform .....	101
Figure 40. Experiment Field View: Right View of the Platform .....	102
Figure 41. Experiment Field Layout: Vertical View. ....	102
Figure 42. Experimental Field Layout: Horizontal view .....	103
Figure 43. True and Estimated Random Trajectory from Experimental Setup. ....	104
Figure 44. Optimal Sensor Configuration .....	104
Figure 45. Position MSE from Experimental Results Demonstrating Waveform Design for PF-TBD ...	105
Figure 46. Joint RF-EO Multi-Target Tracking, 4 Separate Targets in 2-D space .....	105
Figure 47. MSE (Target 1) for MTT with an Increasing Number of RF-EO Sensors .....	106
Figure 48. Measurement with Three Targets .....	107

Figure 49. Mode Probability for 3 Targets at 9 dB Peak SNR .....	108
Figure 50. Target Presence Probability for 3 Targets at 9 dB Peak SNR .....	108
Figure 51. True and Estimated Target Trajectories of 3 Targets at 9 dB Peak SNR .....	109
Figure 52. Target Trajectories for Simulated Target 1 and Target 2 .....	109
Figure 53. Computed MSE for Target 1 for Same and Different Transmit Waveforms .....	110
Figure 54. Computed MSE for Target 2 .....	111
Figure 55. Notional Block Diagram of the Multimodal Radar .....	115
Figure 56. Components and Equipment for the Multimodal Radar .....	116
Figure 57. Field Measurement Set Up for Multimodal Radar .....	117
Figure 58. Flowchart for Operation of Multimodal Radar .....	118
Figure 59. Operation of Multimodal Radar when Scanning a Field .....	119
Figure 60. Imaging Results for a Simulated Target Scenario 1 .....	121
Figure 61. Dual Function for Target Scenario 2 .....	124
Figure 62. Dual Function for Target Scenario 3 .....	125
Figure 63. Dual Function for Target Scenario 4 .....	126
Figure 64. Dependence Of Bandwidth Required for Resolution on Target Strength Ratio .....	127
Figure 65. Scattering Centers of an Extended Target. ....	128
Figure 66. ROC of Multimodal Radar. ....	129
Figure 67. ROC When Multimodal Radar has LRR Pass with 80 MHz Bandwidth .....	130
Figure 68. ROC When Targets are Varying In Strength .....	130
Figure 69. Variation In Required Number of Passes as a Function of Target Separation .....	131
Figure 70. Variation in Required Number of Passes as a Function of SNR .....	132
Figure 71. Membership Functions .....	133
Figure 72. Target Space Divided into Sectors .....	134
Figure 73. Target Scenario 5 .....	134
Figure 74. Multi-Objective Optimization Solution for Sector 3 .....	136
Figure 75. Multi-Objective Optimization Solution for the Problem Scenario .....	136
Figure 76. Scheduling Algorithm for Scenarios with Targets Concentrated in Center .....	137
Figure 77. Scheduling Results with 3 Partial Passes Between Complete Passes .....	138
Figure 78. Tracking Results for Sector 8 with 3 Partial Passes Between Complete Passes. ....	138
Figure 79. Scheduling Results with Complete Passes Only .....	139
Figure 80. Tracking Results with Complete Passes Only (Every Sector) .....	139
Figure 81. Scheduling for Scenarios with Targets Concentrated at the Edges .....	140
Figure 82. Imaging Results for Target Scenario 6 .....	141
Figure 83. Imaging Results for Target Scenario 7 .....	142
Figure 84. Imaging Results for Target Scenario 8 .....	143
Figure 85. Imaging Results for Target Scenario 9 .....	145
Figure 86. Imaging Results for Target Scenario 10 .....	147
Figure 87. HRR Images for Target Scenario 11 .....	149
Figure 88. HRR Images for Target Scenario 12 .....	151

## List of Tables

<b><u>Table</u></b>	<b><u>Page</u></b>
Table 1. Comparison of Our Test Statistic and the Clairvoyant GLRT .....	29
Table 2. Comparison of Our Test Statistic and the Estimated MAP Classifier .....	33
Table 3. Parameters of LFM Chirp Library .....	97
Table 4. Platform Radar System Parameters .....	101
Table 5. Parameters of Waveform Library .....	101
Table 6. Averaged Estimation Position Error .....	103
Table 7. Radar System Parameters .....	116
Table 8. Bandwidth and Resolution for Each Pass of the Multimodal Radar.....	117
Table 9. Target Scenario 1 .....	120
Table 10. Target Scenario 2.....	124
Table 11. Target Scenario 3.....	124
Table 12. Target Scenario 4.....	125
Table 13. Target Scenario 5.....	134
Table 14. Communications Priority .....	135
Table 15. Computed Radar Priority .....	135
Table 16. Simulation Parameters .....	137
Table 17. Target Scenario 6.....	140
Table 18. Target Scenario 7.....	141
Table 19. Target Scenario 8.....	143
Table 20. Target Scenario 9.....	144
Table 21. Target Scenario 10.....	146
Table 22. Target Scenario 11.....	148
Table 23. Target Scenario 12.....	150

## 1. Summary

The Integrated Multi-Modal RF Sensing (IMMRF) project under the Multi-Yield Radio Frequency Countermeasures Investigations and Development (MYRIAD) contract for the Air Force Research Laboratory (AFRL) Sensors Directorate performed basic research to examine key parameters of integrated multi-modal RF sensor design and algorithm development. Four Principal Investigators (PI) representing four different universities executed the program, with each PI being responsible for distinct but interrelated research areas. Dr. Muralidhar Rangaswamy, Principal Electronics Engineer and Technical Advisor for the Radio Frequency Exploitation Branch of the AFRL Sensors Directorate (AFRL/RYP) provided technical direction and collaboration throughout the project.

Dr. Steven Kay from the University of Rhode Island was responsible for researching the *Statistical Framework* for the overall IMMRF project. *Algorithm Development and Implementation* research was conducted by Dr. Lee Swindlehurst from the University of California, Irvine. Dr. Antonia Papandreou-Suppappola from the Arizona State University investigated *Tracking and Waveform Agility*, and Dr. Ram Narayanan from the Pennsylvania State University led the *Radar Test-Bed Development* efforts. Summaries of the results of these four research areas are presented below.

In order to formulate a *Statistical Framework* for multimodal signal processing and distributed systems, we investigated several cases where the joint PDF is not completely known and propose the exponentially embedded family (EEF) to construct the unknown PDF. To obtain optimal performance, we require the joint PDF of the measurements from the sensors, which is not always available. Since the EEF is within the exponential family, it inherits many nice properties of the exponential family, and combines all the available information of the sensors from a statistical standpoint. Also, the maximum likelihood estimation is a convex optimization problem for the EEF due to the convexity of the cumulant generating function, which allows computationally efficient implementation of the EEF.

The *Algorithm Development and Implementation* portion of the research focused on the problem of using multiple, spatially distributed, adaptive multi-modal sensors for multiple target tracking and data association. When left unmodeled or ignored, inherent field-of-view (FOV) limitations for each mode and sensor present a challenge to accurate tracking, appropriate sensor management, and system robustness. We propose a relatively simple variance-penalty oriented modeling solution that effectively presents the FOV as a new design parameter. A novel closed-loop adaptive mode-parameter selection algorithm incorporating this penalty model is also proposed for use in both ideal and cluttered sensing environments.

Under the *Tracking and Waveform Agility* portion of the IMMRF project, we investigated the asymmetrical multi-modal sensing system for tracking low observable targets using radio frequency (RF) radar and electro-optical (EO) sensors. We developed the particle filter (PF) based recursive track before detect (TBD) algorithm for joint RF-EO multi-modal tracking to avoid loss of information caused by matched filter thresholding at low SNR. We also integrated the TBD approach with waveform-agile sensing as the TBD can directly incorporate the transmit waveform in its formulation. The waveform agility is achieved by optimally choosing waveform



parameters to minimize the predicted mean-squared error (MSE) of estimation in tracking. We integrated the embedded exponential family (EEF) approach into the TBD algorithm in order to approximate joint distributions of dependent RF-EO measurements to further improve tracking performance. The waveform-agile multi-modal experimental radar test-bed developed for this project was used to evaluate the proposed algorithms. Finally, we extended the TBD tracker to low observable multiple targets.

For the ***Radar Test-Bed Development***, we developed an adaptive multimodal radar sensor that is capable of progressively varying its range resolution depending upon the target scattering features. Low range resolution profiles are formed using a low bandwidth waveform. High range resolution processing is then performed on selected range cells in which targets are declared. Thus the multimodal radar has the ability to provide target indication with a large range extent and can progressively switch to a narrow range extent mode for extracting recognizable target features. The multimodal radar system developed consists of a test-bed that enables the generation of linear frequency modulated waveforms of various bandwidths for achieving the optimum resolution to image the target. Since bandwidth is a precious resource which can be used by multiple applications, the multimodal radar uses variable bandwidth, making it possible to share the remaining bandwidth with some other application. We explore bandwidth sharing scenarios between radar and communications by dividing the surveillance space into sectors and using fuzzy logic to arrive at priorities for each sector. Finally, we consider the problem of scheduling between tracking and surveillance for the multimodal radar.

## 2. Introduction

In conjunction with the AFOSR sponsored Discovery Challenge Thrust (DCT) topic 1A “Integrated Multi-Modal Sensing, Processing, and Exploitation” in AFOSR-BAA-2007-08, the Air Force Research Laboratory Sensors Directorate, Radio Frequency Exploitation Branch (AFRL/RYP), under the direction of Dr. Muralidhar Rangaswamy, conducted research in Integrated Multi-Modal RF Sensing to support advanced radar signal processing and advanced algorithm development. The Integrated Multi-Modal RF Sensing (IMMRF) project was executed under the Multi-Yield Radio Frequency Countermeasures Investigations and Development (MYRIAD) contract.

The objective of the IMMRF Sensing effort was to perform Basic Research to examine key parameters of integrated multi-modal sensor design and algorithm development. A statistical framework for the program was developed, which ensured the overall statistical validity of the approaches and allowed all researchers to operate from a common probabilistic viewpoint. Dynamic waveform design for agile RF sensing, enhanced detection capabilities, and optimized tracking performance was investigated. Approaches were developed for using multiple, spatially distributed, adaptive multi-modal sensors for multiple target tracking and data association, and a radar test bed was developed to support testing and refinement of theories and algorithms.

Results for each of the four main research areas are presented in the following sections. Detailed discussions of each research area are presented in Sections 3, 4, 5, and 6. Section 3 presents the results of the Statistical Framework research, followed by a discussion of results in the area of Algorithm Development and Implementation in Section 4. Tracking and Waveform Agility research topics are addressed in Section 5, while a detailed presentation of the Radar Test-Bed Development portion of the project is given in Section 6. Conclusions and References are included at the end of the report.

Four Principal Investigators (PI) representing four different universities executed the program, with each PI being responsible for distinct but interrelated research areas.

Dr. Muralidhar Rangaswamy, Principal Electronics Engineer and Technical Advisor for the Radio Frequency Exploitation Branch of the AFRL Sensors Directorate (AFRL/RYP) provided technical direction and collaboration throughout the project.

Dr. Steven Kay from the University of Rhode Island was responsible for researching the **Statistical Framework** for the overall IMMRF project. **Algorithm Development and Implementation** research was conducted by Dr. Lee Swindlehurst from the University of California, Irvine. Dr. Antonia Papandreou-Suppappola from the Arizona State University investigated **Tracking and Waveform Agility**, and Dr. Ram Narayanan from the Pennsylvania State University led the **Radar Test-Bed Development** efforts.

Fundamental mathematical and scientific challenges include:

- (1) Development of high fidelity models to account for the combined RF sensor phenomenology.

- (2) Simulation capability for generating representative scenarios.
- (3) Innovative mathematical approaches for joint processing of RF data for the problems of target detection, tracking, and classification.
- (4) Sophisticated optimization techniques to handle ill-posed inverse problems due to the large dimensionality, multi-modality, and multi-functionality requirements (overcoming the statistical curse of dimensionality).
- (5) Advanced signal and data processing methods, which exploit inherent data characteristics to reduce the training data support and computational cost.
- (6) Novel adaptive processing techniques to deal with non-stationary training data underlying non-side-looking collection geometries, which are not native to conventional radar system operations.
- (7) Adaptive processing on both transmit and receive to maximize system performance; instead of fixed mode sensing, make the sensor adapt to its environment.
- (8) Performance analysis with simulated and measured data using a variety of new statistical metrics.
- (9) Validation through an experimental test-bed and hardware recommendations.

The proposed effort addressed the above issues in a systematic manner in the context of joint, multi-modal detection, tracking and classification through the development of novel mathematical and computational methods to overcome the statistical curse of dimensionality.

### 3. Statistical Framework

Multimodal signal processing and distributed systems have many applications such as radar, sonar, CRBN detection and classification, medical diagnosis, weather prediction, and financial analysis. To obtain optimal performance, we require the joint PDF of the measurements from the sensors, which is not always available.

#### 3.1. Methods, Assumptions, and Procedures

In order to formulate a statistical framework for multimodal signal processing and distributed systems, we investigated several cases where the joint PDF is not completely known and propose the exponentially embedded family (EEF) to construct the unknown PDF. To obtain optimal performance, we require the joint PDF of the measurements from the sensors, which is not always available. Since the EEF is within the exponential family, it inherits many nice properties of the exponential family. For example, the measurements from the sensors form a sufficient statistic for the EEF. The EEF combines all the available information of the sensors from a statistical standpoint. Also, the maximum likelihood estimation is a convex optimization problem for the EEF due to the convexity of the cumulant generating function, which allows computationally efficient implementation of the EEF.

First, we consider the case when the joint PDF is unknown but the marginal PDFs are known. This usually happens when the dimensionality of the sample space is high and we do not have enough training samples to have an accurate estimate of the joint PDF. The problem is exacerbated by onerous environmental and systems constraints in radar and sonar applications. This is also recognized as the “curse of dimensionality” in pattern recognition and machine learning. Hence, it is important to efficiently approximate the unknown joint PDF using limited training data. One common approach is to assume that the measurements from different sensors are conditionally independent. This approach has been widely used due to its simplicity, since the joint PDF is then the product of the marginal PDFs. This is also known as the “product rule” in combining classifiers. In spite of its popularity, the independence assumption may not be a good one if the measurements are actually correlated. Hence researchers have studied other methods that consider the correlation among the measurements. However, the problem does not have a unique solution when the data is non-Gaussian. We propose the EEF that uses the marginal PDFs to estimate the joint PDF that is asymptotically closest to the true one in Kullback-Leibler (KL) divergence.

Note that the above methods all require the knowledge of marginal PDFs. Furthermore, we consider the case when the marginal PDFs are not available or accurate, which can happen due to a high-dimensional sample space and insufficient training data. We present a new way of constructing the joint PDF without the knowledge of marginal PDFs but only a reference PDF. In our method, this reference PDF is the PDF under the null hypothesis, and we assume it is completely known. The constructed joint PDF takes the form of the exponential family and incorporates all the available information. Based on *moment matching*, the parameters in the constructed joint PDF are equivalent to the maximum likelihood estimator (MLE) of the unknown parameters of the exponential family. Hence they can be easily found via convex optimization based on the properties of the exponential family. Since there is no Gaussian

distribution assumption on the reference PDF, this method can be very useful when the underlying distributions are non-Gaussian.

Finally, the EEF is applied to robust signal detection. We examine the case when the alternative hypothesis has multiple candidate models, and apply the multimodal sensor integration technique based on the EEF to detection. It is shown that the EEF is asymptotically optimal as it converges to the true underlying model.

For all the different cases above, simulation results show that the EEF requires less information compared to existing methods but attains comparable performance.

### 3.2. Exponentially Embedded Families for Multimodal Sensor Processing

The exponential embedding of two or more probability density functions is proposed for multimodal sensor processing. It approximates the unknown PDF by exponentially embedding the known PDFs. Such embedding is of an exponential family indexed by some parameters, and hence inherits many nice properties of the exponential family. It is shown that the approximated PDF is asymptotically the one that is the closest to the unknown PDF in Kullback-Leibler (KL) divergence. Applied to hypothesis testing, this approach shows improved performance compared to existing methods for cases of practical importance where the sensor outputs are not independent.

#### 3.2.1. Introduction

Distributed detection systems have many applications such as radar and sonar, medical diagnosis, weather prediction, and financial analysis. To obtain optimal performance, we require the joint PDF of the sensor outputs, which is not always available. One common approach, discussed in [1], [2], is to assume that the PDFs of the sensor outputs are independent, and hence the joint PDF is the product of the marginal PDFs. However, this assumption may not be satisfied since the sensor measurements could be correlated due to the common source and the relative sensor locations. The correlation is noticed in [3], [4], where a copula based framework is proposed to estimate the joint PDF from the marginal PDFs. In this work, we approximate the joint PDF by exponentially embedded families (EEFs) in the sense that it asymptotically minimizes the KL divergence of the true PDF and the estimated one. For two PDFs  $p_1$  and  $p_0$ , the KL divergence is defined as

$$D(p_1 \parallel p_0) = \int p_1(\mathbf{x}) \ln \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})} d\mathbf{x} \quad (1)$$

It is always nonnegative and equals zero if and only if  $p_1 = p_0$  almost everywhere. The KL divergence is a measure of the asymptotic performance of binary hypothesis testing by Stein's lemma [5].

The term "exponentially embedded families" follows that in [6], where it is used for model order estimation. The embedded PDF is of an exponential family indexed by one or more parameters, and so has many nice properties of that family. In a differential geometry point of view, the EEF

forms a manifold in log-PDF space. In one-dimensional case, the EEF is the PDF that minimizes  $D(p \| p_0)$  with the constraint that  $D(p \| p_0) - D(p \| p_1) = \theta$  [5], [7]. Here we focus on the problem of binary hypothesis testing. We assume the presence of two sensors. Similar results are obtained for multiple hypothesis testing and multiple sensors.

### 3.2.2. The EEF and Its Properties

Assume that a source produces the underlying samples  $\mathbf{x}$  which are unobservable, and we have two sensors whose outputs are the statistics  $\mathbf{t}_1(\mathbf{x})$  and  $\mathbf{t}_2(\mathbf{x})$  of  $\mathbf{x}$ . Consider the binary hypothesis testing problem where we know the reference PDF,  $p_{\mathbf{x}}(\mathbf{x}; H_0)$ , but not  $p_{\mathbf{x}}(\mathbf{x}; H_1)$ . So we can find the joint PDF  $p_{\mathbf{T}_1, \mathbf{T}_2}(\mathbf{t}_1, \mathbf{t}_2; H_0)$ , but do not know  $p_{\mathbf{T}_1, \mathbf{T}_2}(\mathbf{t}_1, \mathbf{t}_2; H_1)$ . We assume that the marginal PDFs  $p_{\mathbf{T}_1}(\mathbf{t}_1; H_1)$  and  $p_{\mathbf{T}_2}(\mathbf{t}_2; H_1)$  are known. So the problem is to test between  $H_0$  and  $H_1$  where we know the joint PDF under  $H_0$  and the marginal PDFs under  $H_1$ . The EEF is defined as

$$p_{\mathbf{x}}(\mathbf{x}; \boldsymbol{\eta}) = \frac{\left( \frac{p_{\mathbf{T}_1}(\mathbf{t}_1(\mathbf{x}); H_1)}{p_{\mathbf{T}_1}(\mathbf{t}_1(\mathbf{x}); H_0)} \right)^{\eta_1} \left( \frac{p_{\mathbf{T}_2}(\mathbf{t}_2(\mathbf{x}); H_1)}{p_{\mathbf{T}_2}(\mathbf{t}_2(\mathbf{x}); H_0)} \right)^{\eta_2} p_{\mathbf{x}}(\mathbf{x}; H_0)}{\int \left( \frac{p_{\mathbf{T}_1}(\mathbf{t}_1(\mathbf{x}); H_1)}{p_{\mathbf{T}_1}(\mathbf{t}_1(\mathbf{x}); H_0)} \right)^{\eta_1} \left( \frac{p_{\mathbf{T}_2}(\mathbf{t}_2(\mathbf{x}); H_1)}{p_{\mathbf{T}_2}(\mathbf{t}_2(\mathbf{x}); H_0)} \right)^{\eta_2} p_{\mathbf{x}}(\mathbf{x}; H_0) d\mathbf{x}} \quad (2)$$

where  $\boldsymbol{\eta} = [\eta_1, \eta_2]^T$  are the embedding parameters with the constraints

$$\boldsymbol{\eta} \in \{ \boldsymbol{\eta} : \eta_1, \eta_2 \geq 0, \eta_1 + \eta_2 \leq 1 \} = S \quad (3)$$

Notice that  $p_{\mathbf{x}}(\mathbf{x}; \boldsymbol{\eta})$  does not require the knowledge of  $p_{\mathbf{x}}(\mathbf{x}; H_1)$ . So in practice, we just need to estimate  $p_{\mathbf{x}}(\mathbf{x}; H_0)$  and only the PDFs of  $\mathbf{T}_1$  and  $\mathbf{T}_2$  under  $H_1$  from training data (see also [8]). The reason why we have the constraints in (3) will be explained later. The next theorem is an extension of Kullback's results [5], [7].

**Theorem** *The PDF of  $\mathbf{x}$  as in (2) is the one that minimizes  $D(p_{\mathbf{x}}(\mathbf{x}) \| p_{\mathbf{x}}(\mathbf{x}; H_0))$  subject to the constraints that*

$$D(p_{\mathbf{T}_i}(\mathbf{t}_i) \| p_{\mathbf{T}_i}(\mathbf{t}_i; H_0)) - D(p_{\mathbf{T}_i}(\mathbf{t}_i) \| p_{\mathbf{T}_i}(\mathbf{t}_i; H_1)) = \theta_i \quad (4)$$

for  $i=1, 2$ , where  $p_{\mathbf{T}_1}(\mathbf{t}_1)$  and  $p_{\mathbf{T}_2}(\mathbf{t}_2)$  are the PDFs of  $\mathbf{T}_1$  and  $\mathbf{T}_2$  corresponding to  $p_{\mathbf{x}}(\mathbf{x})$ .

*Proof.* Since

$$D(p_{\mathbf{T}_i}(\mathbf{t}_i) \| p_{\mathbf{T}_i}(\mathbf{t}_i; H_0)) - D(p_{\mathbf{T}_i}(\mathbf{t}_i) \| p_{\mathbf{T}_i}(\mathbf{t}_i; H_1)) = \int p_{\mathbf{x}}(\mathbf{x}) \ln \frac{p_{\mathbf{T}_i}(\mathbf{t}_i(\mathbf{x}); H_1)}{p_{\mathbf{T}_i}(\mathbf{t}_i(\mathbf{x}); H_0)} d\mathbf{x} \quad \text{for } i=1, 2 \quad (5)$$

using Lagrange multipliers for the minimization gives

$$J(p_{\mathbf{x}}(\mathbf{x})) = \int p_{\mathbf{x}}(\mathbf{x}) \ln \frac{p_{\mathbf{x}}(\mathbf{x})}{p_{\mathbf{x}}(\mathbf{x}; H_0)} d\mathbf{x} + \lambda_1 \int p_{\mathbf{x}}(\mathbf{x}) \ln \frac{p_{\mathbf{T}_1}(\mathbf{t}_1(\mathbf{x}); H_1)}{p_{\mathbf{T}_1}(\mathbf{t}_1(\mathbf{x}); H_0)} d\mathbf{x} + \lambda_2 \int p_{\mathbf{x}}(\mathbf{x}) \ln \frac{p_{\mathbf{T}_2}(\mathbf{t}_2(\mathbf{x}); H_1)}{p_{\mathbf{T}_2}(\mathbf{t}_2(\mathbf{x}); H_0)} d\mathbf{x} + \lambda_3 \int p_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \quad (6)$$

Differentiating with respect to  $p_{\mathbf{x}}(\mathbf{x})$  and setting to 0, we have

$$\ln \frac{p_{\mathbf{x}}(\mathbf{x})}{p_{\mathbf{x}}(\mathbf{x}; H_0)} + 1 + \lambda_1 \ln \frac{p_{\mathbf{T}_1}(\mathbf{t}_1(\mathbf{x}); H_1)}{p_{\mathbf{T}_1}(\mathbf{t}_1(\mathbf{x}); H_0)} + \lambda_2 \ln \frac{p_{\mathbf{T}_2}(\mathbf{t}_2(\mathbf{x}); H_1)}{p_{\mathbf{T}_2}(\mathbf{t}_2(\mathbf{x}); H_0)} + \lambda_3 = 0 \quad (7)$$

Solving this equation and letting  $\eta_1 = -\lambda_1$  and  $\eta_2 = -\lambda_2$ , the  $p_{\mathbf{x}}(\mathbf{x})$  that minimizes  $D(p_{\mathbf{x}}(\mathbf{x}) \| p_{\mathbf{x}}(\mathbf{x}; H_0))$  is of the form as in (2) where  $\eta_1$  and  $\eta_2$  are chosen to meet the constraints.

By letting

$$K(\eta) = \ln \int \left( \frac{p_{\mathbf{T}_1}(\mathbf{t}_1(\mathbf{x}); H_1)}{p_{\mathbf{T}_1}(\mathbf{t}_1(\mathbf{x}); H_0)} \right)^{\eta_1} \left( \frac{p_{\mathbf{T}_2}(\mathbf{t}_2(\mathbf{x}); H_1)}{p_{\mathbf{T}_2}(\mathbf{t}_2(\mathbf{x}); H_0)} \right)^{\eta_2} p_{\mathbf{x}}(\mathbf{x}; H_0) d\mathbf{x} \quad (8)$$

$$l_{\mathbf{T}_1}(\mathbf{x}) = \ln \frac{p_{\mathbf{T}_1}(\mathbf{t}_1(\mathbf{x}); H_1)}{p_{\mathbf{T}_1}(\mathbf{t}_1(\mathbf{x}); H_0)}, l_{\mathbf{T}_2}(\mathbf{x}) = \ln \frac{p_{\mathbf{T}_2}(\mathbf{t}_2(\mathbf{x}); H_1)}{p_{\mathbf{T}_2}(\mathbf{t}_2(\mathbf{x}); H_0)} \quad (9)$$

Equation (2) can be written as

$$p_{\mathbf{x}}(\mathbf{x}; \eta) = \exp[\eta_1 l_{\mathbf{T}_1}(\mathbf{x}) + \eta_2 l_{\mathbf{T}_2}(\mathbf{x}) - K(\eta) + \ln p_{\mathbf{x}}(\mathbf{x}; H_0)] \quad (10)$$

which is a two-parameter exponential family [9].  $K(\eta)$  is recognized as the cumulant generating function of  $l_{\mathbf{T}_1}(\mathbf{x})$ ,  $l_{\mathbf{T}_2}(\mathbf{x})$  when the PDF of  $\mathbf{x}$  is  $p_{\mathbf{x}}(\mathbf{x}; H_0)$ . Since (10) is of an exponential family, the EEF inherits some useful properties that we will discuss in the following (refer to [9], [10] and [11] for details).

1) If the PDF of  $\mathbf{x}$  is  $p_{\mathbf{x}}(\mathbf{x}; \eta)$ , then the joint PDF of  $\mathbf{T}_1$  and  $\mathbf{T}_2$  is [11]

$$p_{\mathbf{T}_1, \mathbf{T}_2}(\mathbf{t}_1, \mathbf{t}_2; \eta) = \exp[\eta_1 l_{\mathbf{T}_1} + \eta_2 l_{\mathbf{T}_2} - K(\eta) + \ln p_{\mathbf{T}_1, \mathbf{T}_2}(\mathbf{t}_1, \mathbf{t}_2; H_0)] \quad (11)$$

where

$$l_{\mathbf{T}_1} = \ln \frac{p_{\mathbf{T}_1}(\mathbf{t}_1; H_1)}{p_{\mathbf{T}_1}(\mathbf{t}_1; H_0)}, l_{\mathbf{T}_2} = \ln \frac{p_{\mathbf{T}_2}(\mathbf{t}_2; H_1)}{p_{\mathbf{T}_2}(\mathbf{t}_2; H_0)} \quad (12)$$

This can also be easily proved using surface integral techniques [12]. Notice that for (11),  $\mathbf{T}_1$  and  $\mathbf{T}_2$  are not independent unless they are independent under  $H_0$ .

2)  $K(\eta)$  is convex by Holder's inequality [9]. If we assume that  $l_{\mathbf{T}_1}$  and  $l_{\mathbf{T}_2}$  are linearly independent [13], then  $\eta$  is identifiable, and hence  $K(\eta)$  is strictly convex [10].

3) Let  $E_\eta(l_{\mathbf{T}_i})$  be the expected value of  $l_{\mathbf{T}_i}$  for  $i=1,2$  and  $\mathbf{C}_\eta$  be the covariance matrix of  $[l_{\mathbf{T}_1}, l_{\mathbf{T}_2}]^T$  when  $\mathbf{x}$  is distributed according to  $p_{\mathbf{x}}(\mathbf{x}; \eta)$ . We have

$$\frac{\partial K(\eta)}{\partial \eta_i} = E_\eta(l_{\mathbf{T}_i}) \quad (13)$$

$$\begin{bmatrix} \frac{\partial^2 K(\eta)}{\partial \eta_1^2} & \frac{\partial^2 K(\eta)}{\partial \eta_1 \partial \eta_2} \\ \frac{\partial^2 K(\eta)}{\partial \eta_2 \partial \eta_1} & \frac{\partial^2 K(\eta)}{\partial \eta_2^2} \end{bmatrix} = \mathbf{C}_\eta \quad (14)$$

Notice that (14) also shows that  $K(\eta)$  is convex.

4)  $[l_{\mathbf{T}_1}, l_{\mathbf{T}_2}]^T$  is a minimal and complete sufficient statistic for  $\eta$ . Hence  $[l_{\mathbf{T}_1}, l_{\mathbf{T}_2}]^T$  can be used to discriminate between  $p_{\mathbf{x}}(\mathbf{x}; H_1)$  and  $p_{\mathbf{x}}(\mathbf{x}; H_0)$ .

5)  $K(\eta)$  is finite on  $S$ . To see this,  $K(\eta) > -\infty$  by definition. Obviously,  $K(\eta) = 0$  for  $\eta = [0,0]^T, [1,0]^T, [0,1]^T$ . Since  $K(\eta)$  is strictly convex, we have  $K(\eta) \leq 0 < \infty$  for  $\eta \in S$ . But when  $\eta$  is outside  $S$ , there is no guarantee that  $K(\eta)$  is finite in general. This explains why we have the constraints in (3).

### 3.2.3. EEf for Hypothesis Testing

For binary hypothesis testing, we will decide  $H_1$  if



$$\max_{\eta} \ln \frac{p_{\mathbf{x}}(\mathbf{x}; \eta)}{p_{\mathbf{x}}(\mathbf{x}; H_0)} > \tau \quad (15)$$

where  $\tau$  is a threshold. This test statistic actually does not depend on  $\mathbf{x}$  but only on  $\mathbf{t}_1$  and  $\mathbf{t}_2$  since

$$g(\eta) = \ln \frac{p_{\mathbf{x}}(\mathbf{x}; \eta)}{p_{\mathbf{x}}(\mathbf{x}; H_0)} = \eta_1 l_{\mathbf{T}_1} + \eta_2 l_{\mathbf{T}_2} - K(\eta) \quad (16)$$

The reason why we choose this test statistic, as we will show next, is that asymptotically  $\max_{\eta} p_{\mathbf{x}}(\mathbf{x}; \eta)$  is the closest to the unknown  $p_{\mathbf{x}}(\mathbf{x}; H_1)$  in KL divergence.

Assume that there are a large number of independent and identically distributed (IID) unobservable  $\mathbf{x}_i$ 's for  $i=1, 2, \dots, N$ , which results in IID  $\mathbf{t}_{1i}$ 's and IID  $\mathbf{t}_{2i}$ 's. We want to maximize

$$\frac{1}{N} \sum_{i=1}^N \ln \frac{p_{\mathbf{x}}(\mathbf{x}_i; \eta)}{p_{\mathbf{x}}(\mathbf{x}_i; H_0)} = \exp \left[ \eta_1 \frac{1}{N} \sum_{i=1}^N l_{\mathbf{T}_{1i}} + \eta_2 \frac{1}{N} \sum_{i=1}^N l_{\mathbf{T}_{2i}} - K(\eta) \right] \quad (17)$$

By the law of large numbers, under  $H_1$

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N l_{\mathbf{T}_{1i}} &\rightarrow E_{H_1} \left[ l_{\mathbf{T}_1} \right] = D \left( p_{\mathbf{T}_1}(\mathbf{t}_1; H_1) \parallel p_{\mathbf{T}_1}(\mathbf{t}_1; H_0) \right) \\ \frac{1}{N} \sum_{i=1}^N l_{\mathbf{T}_{2i}} &\rightarrow E_{H_1} \left[ l_{\mathbf{T}_2} \right] = D \left( p_{\mathbf{T}_2}(\mathbf{t}_2; H_1) \parallel p_{\mathbf{T}_2}(\mathbf{t}_2; H_0) \right) \end{aligned} \quad (18)$$

as  $N \rightarrow \infty$ . So we are asymptotically maximizing

$$\eta_1 D \left( p_{\mathbf{T}_1}(\mathbf{t}_1; H_1) \parallel p_{\mathbf{T}_1}(\mathbf{t}_1; H_0) \right) + \eta_2 D \left( p_{\mathbf{T}_2}(\mathbf{t}_2; H_1) \parallel p_{\mathbf{T}_2}(\mathbf{t}_2; H_0) \right) - K(\eta) \quad (19)$$

Since

$$\ln \frac{p_{\mathbf{x}}(\mathbf{x}; H_1)}{p_{\mathbf{x}}(\mathbf{x}; \eta)} = -\eta_1 l_{\mathbf{T}_1} - \eta_2 l_{\mathbf{T}_2} + K(\eta) + \ln \frac{p_{\mathbf{x}}(\mathbf{x}; H_1)}{p_{\mathbf{x}}(\mathbf{x}; H_0)} \quad (20)$$

the KL divergence between  $p_{\mathbf{x}}(\mathbf{x}; H_1)$  and  $p_{\mathbf{x}}(\mathbf{x}; \eta)$  is

$$\begin{aligned}
& D(p_{\mathbf{x}}(\mathbf{x}; H_1) \| p_{\mathbf{x}}(\mathbf{x}; \eta)) \\
&= E_{H_1} \exp \left[ -\eta_1 l_{\mathbf{T}_1} - \eta_2 l_{\mathbf{T}_2} + K(\eta) + \ln \frac{p_{\mathbf{x}}(\mathbf{x}; H_1)}{p_{\mathbf{x}}(\mathbf{x}; H_0)} \right] \\
&= -\eta_1 D(p_{\mathbf{T}_1}(\mathbf{t}_1; H_1) \| p_{\mathbf{T}_1}(\mathbf{t}_1; H_0)) - \eta_2 D(p_{\mathbf{T}_2}(\mathbf{t}_2; H_1) \| p_{\mathbf{T}_2}(\mathbf{t}_2; H_0)) \\
&\quad + K(\eta) + D(p_{\mathbf{x}}(\mathbf{x}; H_1) \| p_{\mathbf{x}}(\mathbf{x}; H_0))
\end{aligned} \tag{21}$$

This shows that  $D(p_{\mathbf{x}}(\mathbf{x}; H_1) \| p_{\mathbf{x}}(\mathbf{x}; \eta))$  is minimized by maximizing (19). A similar result is shown in [6] by using a Pythagorean-like theorem. Also if  $\mathbf{T}_1$  and/or  $\mathbf{T}_2$  are sufficient statistics for deciding between  $H_0$  and  $H_1$ , it can be shown that  $p_{\mathbf{x}}(\mathbf{x}; \eta) = p_{\mathbf{x}}(\mathbf{x}; H_1)$ . Thus, the true PDF under  $H_1$  is recovered [14].

To implement (15), we require the maximum likelihood estimate (MLE) of  $\eta$ . Let  $\eta^*$  be the MLE of  $\eta$  without constraints in (3). Since  $g(\eta)$  is strictly concave,  $\eta^*$  is unique. Taking partial derivatives of  $g(\eta)$  and setting to 0, we have

$$l_{\mathbf{T}_1} = \frac{\partial K(\boldsymbol{\eta})}{\partial \eta_1} \Big|_{\boldsymbol{\eta}^*}, l_{\mathbf{T}_2} = \frac{\partial K(\boldsymbol{\eta})}{\partial \eta_2} \Big|_{\boldsymbol{\eta}^*} \tag{22}$$

Let  $\eta$  be the MLE of  $\eta$  with the constraints. If  $\eta^*$  is in the constraint set  $S$ , then  $\eta = \eta^*$ . Otherwise,  $\eta$  is unique and is on the boundary of  $S$  since  $-g(\eta)$  is strictly convex and  $S$  is convex also [15], and hence we could simply search the boundary of  $S$  to find  $\eta$ .

### 3.2.4. Example

Since only  $\mathbf{T}_1$  and  $\mathbf{T}_2$  are used in hypothesis testing, we only need to specify their distributions. Consider the case when  $\mathbf{T}_1$  and  $\mathbf{T}_2$  are scalars (will write them as  $T_1$  and  $T_2$ ) with distributions

$$\begin{bmatrix} T_1 \\ T_2 \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma^2 \begin{bmatrix} 1 & \rho_0 \\ \rho_0 & 1 \end{bmatrix} \right) \quad \text{under } H_0 \tag{23}$$

$$\begin{bmatrix} T_1 \\ T_2 \end{bmatrix} \sim N \left( \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, \sigma^2 \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix} \right) \quad \text{under } H_1 \tag{24}$$

where  $\rho_0$  is known but  $\rho_1$  is unknown (we do not need the joint PDF of  $T_1$  and  $T_2$  under  $H_1$ ).

We have

$$\begin{aligned}
K(\eta) &= \ln E_{H_0} \left[ \exp \left( \eta_1 l_{T_1} + \eta_2 l_{T_2} \right) \right] \\
&= \ln E_{H_0} \left[ \exp \left( \eta_1 \frac{2t_1 A_1 - A_1^2}{2\sigma^2} + \eta_2 \frac{2t_2 A_2 - A_2^2}{2\sigma^2} \right) \right] \\
&= -\eta_1 \frac{A_1^2}{2\sigma^2} - \eta_2 \frac{A_2^2}{2\sigma^2} + \ln E_{H_0} \left[ \exp \left( \frac{\eta_1 A_1 + \eta_2 A_2}{\sigma^2} \right) \right]
\end{aligned} \tag{25}$$

Let  $\phi = \left[ \frac{\eta_1 A_1}{\sigma^2}, \frac{\eta_2 A_2}{\sigma^2} \right]^T$  and  $\mathbf{t} = [t_1, t_2]^T$ , then

$$E_{H_0} \left[ \exp \left( \frac{\eta_1 t_1 A_1 + \eta_2 t_2 A_2}{\sigma^2} \right) \right] = E_{H_0} \left[ \exp \left( \phi^T \mathbf{t} \right) \right] = \exp \left( \frac{1}{2} \phi^T \mathbf{C}_0 \phi \right) \tag{26}$$

where  $\mathbf{C}_0 = \sigma^2 \begin{bmatrix} 1 & \rho_0 \\ \rho_0 & 1 \end{bmatrix}$  and hence

$$K(\eta) = -\eta_1 \frac{A_1^2}{2\sigma^2} - \eta_2 \frac{A_2^2}{2\sigma^2} + \frac{1}{2} \phi^T \mathbf{C}_0 \phi \tag{27}$$

So

$$\begin{aligned}
g(\eta) &= \eta_1 l_{T_1} + \eta_2 l_{T_2} - K(\eta) \\
&= \eta_1 \frac{2t_1 A_1 - A_1^2}{2\sigma^2} + \eta_2 \frac{2t_2 A_2 - A_2^2}{2\sigma^2} - K(\eta) \\
&= \frac{\eta_1 A_1 t_1}{\sigma^2} + \frac{\eta_2 A_2 t_2}{\sigma^2} - \frac{1}{2} \phi^T \mathbf{C}_0 \phi \\
&= \mathbf{t}^T \phi - \frac{1}{2} \phi^T \mathbf{C}_0 \phi
\end{aligned} \tag{28}$$

Differentiating and setting to 0, the global maximum is found at

$$\phi^* = \mathbf{C}_0^{-1} \mathbf{t} = \begin{bmatrix} \frac{t_1 - \rho_0 t_2}{1 - \rho_0^2} \\ \frac{t_2 - \rho_0 t_1}{1 - \rho_0^2} \end{bmatrix} \tag{29}$$

or

$$\eta^* = \begin{bmatrix} \frac{\sigma^2 (t_1 - \rho_0 t_2)}{A_1 (1 - \rho_0^2)} \\ \frac{\sigma^2 (t_2 - \rho_0 t_1)}{A_2 (1 - \rho_0^2)} \end{bmatrix} \tag{30}$$

If  $\eta^* \in S$ , then we decide  $H_1$  if  $g(\eta^*) = \mathbf{t}^T \mathbf{C}_0^{-1} \mathbf{t} > \tau$ , otherwise we search  $\eta$  on the boundary and decide  $H_1$  if  $g(\eta) > \tau$ .

When we observe  $N$  IID  $t_{1i}$ 's and IID  $t_{2i}$ 's, then it easily extends that by (17),  $[t_1, t_2]^T$  is replaced by the sample mean  $[\frac{1}{N} \sum_{i=1}^N t_{1i}, \frac{1}{N} \sum_{i=1}^N t_{2i}]^T$ , and everything else remains the same.

### 3.2.5. Simulation Results

For the above example, we set  $N = 20$ ,  $A_1 = 0.3$ ,  $A_2 = 0.35$ ,  $\sigma^2 = 1$ ,  $\rho_0 = 0.6$  and  $\rho_1 = 0.7$ . We compare the EEF approach with the clairvoyant detector ( $\rho_1$  is known, its performance is an upper bound), the detector assuming independence of  $t_1$  and  $t_2$ , and the copula based method. The copula method estimates the linear correlation coefficient  $\rho_1$  using a non-parametric rank correlation measure, Kendall's  $\tau$ . We use the Gaussian copula as in [3]. The simulation is repeated for 5000 trials. The receiver operating characteristic curves (ROC) are plotted. As seen in Figure 1, the EEF is only poorer than the clairvoyant detector, and performs better than the other two methods.

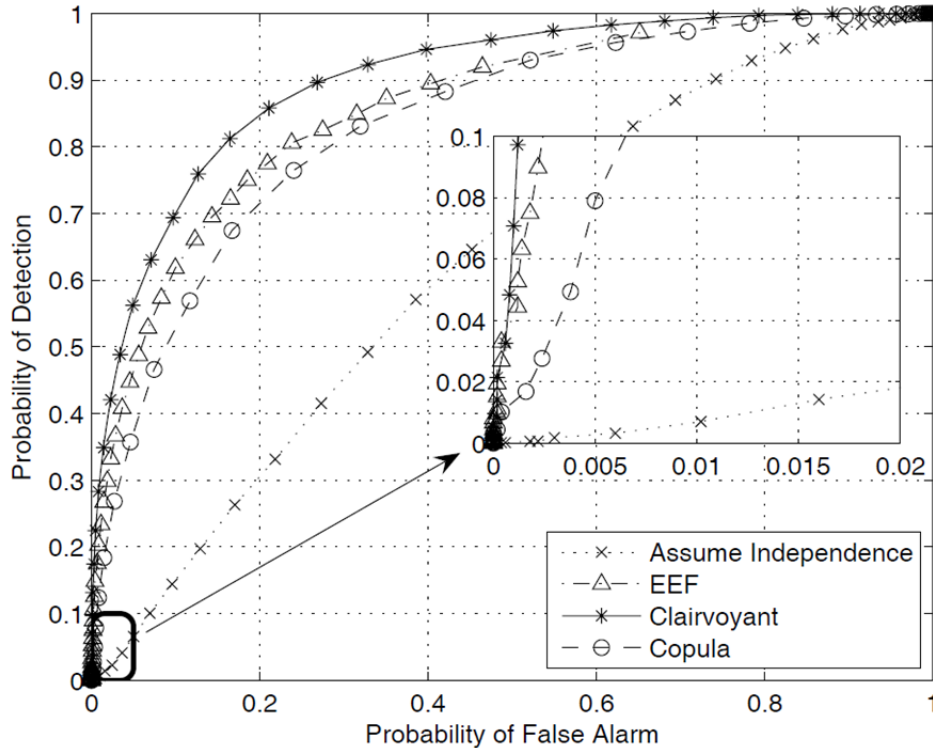


Figure 1. ROC Curves for Different Detectors.

### 3.3. Sensor Integration by Joint PDF Construction Using the EEF

In this section, we investigate the problem of sensor integration to combine all the available information in a multi-sensor setting from a statistical standpoint. Specifically, we propose a novel method of constructing the joint probability density function (PDF) of the measurements from all the sensors based on the exponential family and small signal assumption. The constructed PDF only requires knowledge of the joint PDF under a reference hypothesis and hence is useful in many practical cases. Examples and simulation results show that our method requires less information compared to existing methods but attains comparable detection and classification performance.

#### 3.3.1. Introduction

Distributed systems and information fusion have been widely studied and used in engineering, finance, and scientific research. Such applications are to radar, sonar, biomedical analysis, stock prediction, weather forecasting, and chemical, biological, radiological, and nuclear (CBRN) detection, to name a few. If the joint probability density function (PDF) under each candidate hypothesis is known, we would easily obtain the optimal performance by the Neyman-Pearson rule (from the frequentist inference) for detection (binary hypothesis testing) or by the maximum a posteriori probability (MAP) rule (from the Bayesian inference) for both detection and classification (multiple hypothesis testing) [16]. However in practice, this information may not be available. This usually happens when the dimensionality of the sample space is high and we do not have enough training samples to have an accurate estimate of the joint PDF. The problem is exacerbated by onerous environmental and systems constraints in radar and sonar applications. This is also recognized as the “curse of dimensionality” in pattern recognition and machine learning. Hence, it is important to efficiently approximate the unknown joint PDF using limited training data. One common approach is to assume that the measurements from different sensors are conditionally independent [17], [18]. This approach has been widely used due to its simplicity, since the joint PDF is then the product of the marginal PDFs. This is also known as the “product rule” in combining classifiers [19]. In spite of its popularity, the independence assumption may not be a good one if the measurements are actually correlated. Furthermore, as stated in [19], the product rule is severe because “it is sufficient for a single recognition engine to inhibit a particular interpretation by outputting a close to zero probability for it”. Hence researchers have studied other methods that consider the correlation among the measurements. However, the problem does not have a unique solution when the data is non-Gaussian. A copula based framework is proposed in [20], [21] to construct the joint PDF. The exponentially embedded families (EEFs) are used in [22] to estimate the joint PDF that is asymptotically closest to the true one in Kullback-Leibler (KL) divergence.

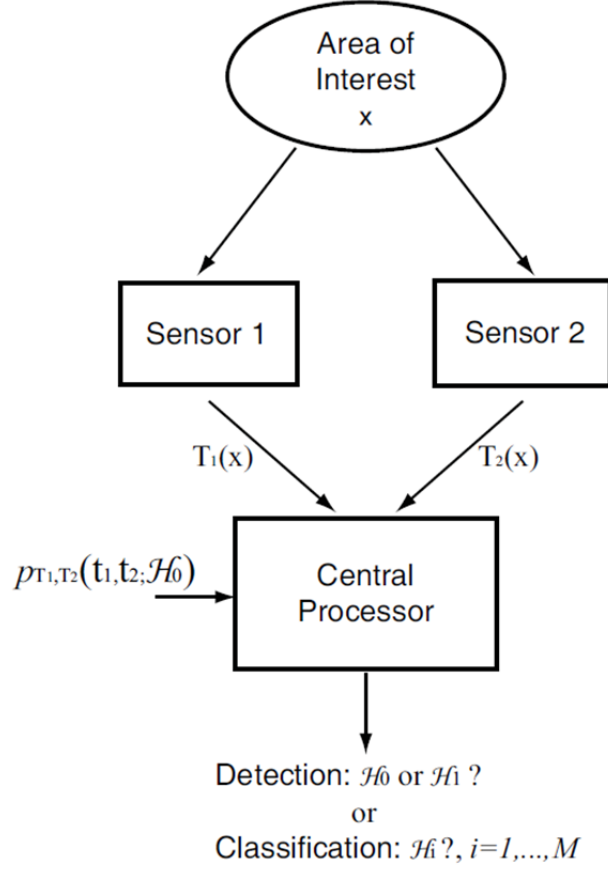
Note that the above methods all require the knowledge of marginal PDFs. In this paper, we consider the case when the marginal PDFs are not available or accurate, which can happen due to a high-dimensional sample space and insufficient training data. We present a new way of constructing the joint PDF without the knowledge of marginal PDFs but only a reference PDF. In our method, this reference PDF is the PDF under the null hypothesis  $H_0$ , and we assume it is completely known. The constructed joint PDF takes the form of the exponential family and incorporates all the available information. Based on *moment matching*, the parameters in the

constructed joint PDF are equivalent to the maximum likelihood estimator (MLE) [23] of the unknown parameters of the exponential family. Hence they can be easily found via convex optimization based on the properties of the exponential family. Since there is no Gaussian distribution assumption on the reference PDF, this method can be very useful when the underlying distributions are non-Gaussian. We start with the detection problem, and then extend our method to the classification problem. For detection, it is shown that under some conditions, our detection statistics are the same as the clairvoyant generalized likelihood ratio test (GLRT). For classification, our classifier also has the same performance as the estimated MAP classifier. Both the clairvoyant GLRT and the estimated MAP classifier assume that the true PDFs under each candidate hypothesis are known except for the usual unknown parameters.

## Problem Statement

Consider the distributed detection/classification problem when we observe the outputs of two sensors,  $\mathbf{T}_1(\mathbf{x})$  and  $\mathbf{T}_2(\mathbf{x})$ , which are transformations of the underlying samples  $\mathbf{x}$ . The latter are unobservable at the central processor as shown in Figure 1. We choose two sensors for simplicity. All the results in this paper are valid for multiple sensors. For detection, we want to distinguish between two hypotheses  $H_0$  and  $H_1$  based on the outputs of the two sensors, and for classification, we have  $M$  candidate hypotheses  $H_i$  for  $i = 1, 2, \dots, M$ .

Assume that we have sufficient training data  $\mathbf{T}_1^{(n)}(\mathbf{x})$ 's and  $\mathbf{T}_2^{(n)}(\mathbf{x})$ 's under  $H_0$ , i.e., when there is no signal present. Hence, we have a good estimate of the joint PDF of  $\mathbf{T}_1$  and  $\mathbf{T}_2$  under  $H_0$  [9], and thus we assume  $p_{\mathbf{T}_1, \mathbf{T}_2}(\mathbf{t}_1, \mathbf{t}_2; H_0)$  is completely known. Under  $H_1$  for detection or  $H_i$  for  $i = 1, 2, \dots, M$  for classification when a signal is present, we may not even have enough training data to estimate the marginal PDFs at each sensor, let alone the joint PDF. This is especially the case in the radar scenario, where the target is present for only a small portion of the time. So our goal is to use the available information to construct an appropriate  $p_{\mathbf{T}_1, \mathbf{T}_2}(\mathbf{t}_1, \mathbf{t}_2; H_1)$  under  $H_1$  for detection or  $p_{\mathbf{T}_1, \mathbf{T}_2}(\mathbf{t}_1, \mathbf{t}_2; H_i)$  under each  $H_i$  for classification. A simple illustration is shown in Figure 2.



**Figure 2. Distributed Detection/Classification System with Two Sensors**

### 3.3.2. Joint PDF Construction by Exponential Family and Its Application in Distributed Systems

First, we consider the detection problem, where we wish to construct  $p_{\mathbf{T},\mathbf{T}_2}(\mathbf{t}_1, \mathbf{t}_2; H_1)$ . The result will then be extended to the classification problem.

To simplify the notation, let  $\mathbf{T} = [\mathbf{T}_1 \ \mathbf{T}_2]^T$  so that we can write the joint PDF  $p_{\mathbf{T},\mathbf{T}_2}(\mathbf{t}_1, \mathbf{t}_2; H_i)$  as  $p_{\mathbf{T}}(\mathbf{t}; H_i)$  for  $i = 0, 1$ .

Since  $p_{\mathbf{T}}(\mathbf{t}; H_1)$  cannot be uniquely specified based on  $p_{\mathbf{T}}(\mathbf{t}; H_0)$ , we will assume that 1)  $p_{\mathbf{T}}(\mathbf{t}; H_1)$  is close to  $p_{\mathbf{T}}(\mathbf{t}; H_0)$  (small signal assumption), and 2) the expected value of  $\mathbf{T}$  under  $H_1$  or  $E_1(\mathbf{T})$  is known. The reason that we assume small signal is because in practice, we are really interested in the small signal case, since for large signals, even a non-optimal detector would have acceptable performance. Now, we want to find a PDF  $\hat{p}_{\mathbf{T}}(\mathbf{t})$  such that the KL divergence

$$D(\hat{p}_{\mathbf{T}}(\mathbf{t}) \| p_{\mathbf{T}}(\mathbf{t}; H_0)) = \int \hat{p}_{\mathbf{T}}(\mathbf{t}) \ln \frac{\hat{p}_{\mathbf{T}}(\mathbf{t})}{p_{\mathbf{T}}(\mathbf{t}; H_0)} d\mathbf{t} \quad (31)$$

is minimized with the constraint that

$$E_{\hat{p}}(\mathbf{T}) = E_1(\mathbf{T}) \quad (32)$$

Then the PDF  $\hat{p}_{\mathbf{T}}(\mathbf{t})$  is used as our constructed PDF of  $\mathbf{T}$  under  $H_1$ . Note that the constraint in (32) is considered as *moment matching*, since the constructed PDF has the same moment as the true PDF. Here we consider the KL divergence because by Stein's lemma [25], the KL divergence determines the asymptotic performance for detection. An extended result to classification has been recently presented in [26]. Therefore, this is a worst-case approach obtained by minimizing  $D(\hat{p}_{\mathbf{T}}(\mathbf{t}) \| p_{\mathbf{T}}(\mathbf{t}; H_0))$ .

The Kullback theorem in [27] shows that the solution of the above problem is

$$\hat{p}_{\mathbf{T}}(\mathbf{t}) = \exp[\boldsymbol{\theta}^T \mathbf{t} - K(\boldsymbol{\theta}) + \ln p_{\mathbf{T}}(\mathbf{t}; H_0)] \quad (33)$$

where

$$K(\boldsymbol{\theta}) = \ln E_0[\exp(\boldsymbol{\theta}^T \mathbf{T})] = \ln \int \exp(\boldsymbol{\theta}^T \mathbf{T}) p_{\mathbf{T}}(\mathbf{t}; H_0) d\mathbf{t} \quad (34)$$

is the cumulant generating function of  $p_{\mathbf{T}}(\mathbf{t}; H_0)$ , and it normalizes the PDF to integrate to one. Here  $\boldsymbol{\theta}$  are not free parameters, and it has to satisfy the constraint in (32). It can be easily shown that

$$\frac{\partial K(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = E_{\hat{p}}(\mathbf{T}) \quad (35)$$

where  $\frac{\partial K(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \left[ \frac{\partial K(\boldsymbol{\theta})}{\partial \theta_1} \frac{\partial K(\boldsymbol{\theta})}{\partial \theta_2} \dots \frac{\partial K(\boldsymbol{\theta})}{\partial \theta_p} \right]^T$  with  $p$  being the length of  $\boldsymbol{\theta}$ . Therefore, the constraint in (32) is equivalent to

$$\frac{\partial K(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = E_1(\mathbf{T}) \quad (36)$$

Note that in practice, we may not know  $E_1(\mathbf{T})$ , and therefore, we use  $\mathbf{t}$  as an estimate of  $E_1(\mathbf{T})$ . Finally, the constructed PDF under  $H_1$  is given by (33) where  $\boldsymbol{\theta}$  satisfies

$$\frac{\partial K(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{t} \quad (37)$$

Now suppose we have an exponential family parameterized by  $\boldsymbol{\theta}$  as



$$p_{\mathbf{T}}(\mathbf{t}; \boldsymbol{\theta}) = \exp \left[ \boldsymbol{\theta}^T \mathbf{t} - K(\boldsymbol{\theta}) + \ln p_{\mathbf{T}}(\mathbf{t}; H_0) \right] \quad (38)$$

which is the same as in (32) except that  $\boldsymbol{\theta}$  are free parameters. It is also shown in [28] that families of PDFs with small statistical curvature enjoy, nearly, the good statistical properties of the EEF in (38). Since  $K(\boldsymbol{\theta})$  is convex by Holder's inequality [29], the MLE of  $\boldsymbol{\theta}$  can be obtained by taking the derivative of  $\boldsymbol{\theta}^T \mathbf{t} - K(\boldsymbol{\theta})$  with respect to  $\boldsymbol{\theta}$  and setting it to zero. This results in

$$\mathbf{t} - \frac{\partial K(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{0} \quad (39)$$

which is identical to (37). Assume that  $K(\boldsymbol{\theta})$  is strictly convex, and the solution is unique. Therefore, this shows that the MLE produces the same  $\boldsymbol{\theta}$  as moment matching does in (37). Hence, we can write the constructed PDF  $\hat{p}_{\mathbf{T}}(\mathbf{t})$  in (33) as

$$\hat{p}_{\mathbf{T}}(\mathbf{t}) = p_{\mathbf{T}}(\mathbf{t}; \hat{\boldsymbol{\theta}}) = \exp \left[ \hat{\boldsymbol{\theta}}^T \mathbf{t} - K(\hat{\boldsymbol{\theta}}) + \ln p_{\mathbf{T}}(\mathbf{t}; H_0) \right] \quad (40)$$

where  $\hat{\boldsymbol{\theta}}$  is the MLE of  $\boldsymbol{\theta}$ . Also note that since  $K(\boldsymbol{\theta})$  is convex, finding the MLE becomes a convex optimization problem and many existing methods can be readily utilized [30], [31]. If  $\mathbf{t} = \mathbf{0}$ , then the PDF in (38) becomes  $p_{\mathbf{T}}(\mathbf{t}; H_0)$  or

$$p_{\mathbf{T}}(\mathbf{t}; \mathbf{0}) = p_{\mathbf{T}}(\mathbf{t}; H_0) \quad (41)$$

This shows that  $p_{\mathbf{T}}(\mathbf{t}; H_0)$  also belongs to the exponential family in (38).

This constructed PDF in (40) looks similar to the Edgeworth expansion (see equations (2.1) and (2.17) in [32]). The Edgeworth expansion is an approximation of the cumulative distribution function (CDF) of the sum of IID samples starting from the CDF of the standard Gaussian distribution. Note that the PDF in (38) belongs to the exponential family. Since  $\mathbf{T}$  is a sufficient statistic for the exponential PDF in (38), this PDF incorporates all the information from the two sensors. Since  $p_{\mathbf{T}}(\mathbf{t}; H_0)$  is required to construct  $p_{\mathbf{T}}(\mathbf{t}; \hat{\boldsymbol{\theta}})$ , and it is assumed that  $p_{\mathbf{T}}(\mathbf{t}; H_0)$  is available or it can be estimated with reasonable accuracy. Also note that if  $\mathbf{T}_1$ ,  $\mathbf{T}_2$  are statistically dependent under  $H_0$ , they will also be dependent for the constructed PDF under  $H_1$ .

With the small signal assumption,  $p_{\mathbf{T}}(\mathbf{t}; H_1)$  is close to  $p_{\mathbf{T}}(\mathbf{t}; H_0)$ . Under this constraint, it has been shown in [33] that if  $\mathbf{t}$  is the score function, i.e.,

$$\mathbf{t} = \frac{\partial \ln p_{\mathbf{T}}(\mathbf{t}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\mathbf{0}} \quad (42)$$

by using a first order Taylor expansion on the log-likelihood function  $\ln p_{\mathbf{T}}(\mathbf{t};)$  about  $\mathbf{0}$ , we obtain the same PDF as in (38). This small signal analysis is similar to the locally optimum detector (LOD) [34], but there are some fundamental differences between our method and the LOD:

- (i) Our method produces a PDF, not only a test statistic.
- (ii)  $\mathbf{t}$  need not be a score function for the constructed PDF to be a valid PDF. If it were, the constructed PDF in (38) could be interpreted as a first-order Taylor expansion and normalized to be a PDF.

Finally, since (40) is the constructed PDF under  $H_1$ , we decide  $H_1$  if

$$\ln \frac{p_{\mathbf{T}}(\mathbf{t}; \hat{\boldsymbol{\theta}})}{p_{\mathbf{T}}(\mathbf{t}; H_0)} = \hat{\boldsymbol{\theta}}^T \mathbf{t} - K(\hat{\boldsymbol{\theta}}) > \tau \quad (43)$$

where  $\tau$  is a threshold. Note that Kullback also had similar ideas (see Chapter 5 in [27] where  $\hat{\boldsymbol{\theta}}^T \mathbf{t} - K(\hat{\boldsymbol{\theta}})$  can be considered as the estimated KL divergence between  $p_{\mathbf{T}}(\mathbf{t}; \boldsymbol{\theta})$  and  $p_{\mathbf{T}}(\mathbf{t}; H_0)$ ). This method can be extended to the classification problem. Similar to (40), as shown in [35], we can construct the PDF of  $\mathbf{T}$  under  $H_i$  as

$$\hat{p}_{\mathbf{T}}(\mathbf{t}; \boldsymbol{\theta}_i) = p_{\mathbf{T}}(\mathbf{t}; \hat{\boldsymbol{\theta}}_i) = \exp \left[ \hat{\boldsymbol{\theta}}_i^T \mathbf{t} - K(\hat{\boldsymbol{\theta}}_i) + \ln p_{\mathbf{T}}(\mathbf{t}; H_0) \right] \quad (44)$$

where

$$K(\boldsymbol{\theta}_i) = \ln E_0 \left[ \exp(\boldsymbol{\theta}_i^T \mathbf{T}) \right] \quad (45)$$

is the cumulant generating function of  $p_{\mathbf{T}}(\mathbf{t}; H_0)$  that normalizes the constructed PDF, and  $p_{\mathbf{T}}(\mathbf{t}; \hat{\boldsymbol{\theta}}_i)$  is considered as our estimate of  $p_{\mathbf{T}}(\mathbf{t}; H_i)$  where  $\hat{\boldsymbol{\theta}}_i$  is the MLE of  $\boldsymbol{\theta}_i$ . Hence similar to the MAP rule [1], we will decide  $H_i$  for which the following is maximum over  $i$ :

$$\frac{\hat{p}(H_i | \mathbf{t})}{p_{\mathbf{T}}(\mathbf{t})} = \hat{p}_{\mathbf{T}}(\mathbf{t}; H_i) p(H_i) = p_{\mathbf{T}}(\mathbf{t}; \hat{\boldsymbol{\theta}}_i) p(H_i) \quad (46)$$

When we assume that the prior probabilities of each candidate hypothesis are equal, i.e.,  $p(H_1) = \dots = p(H_M) = 1/M$ ,  $p(H_i)$  cancels and we can equivalently decide  $H_i$  for which the following is maximum over  $i$ :

$$\ln \frac{p_{\mathbf{T}}(\mathbf{t}; \hat{\boldsymbol{\theta}}_i)}{p_{\mathbf{T}}(\mathbf{t}; H_0)} = \hat{\boldsymbol{\theta}}_i^T \mathbf{t} - K(\hat{\boldsymbol{\theta}}_i) \quad (47)$$

### 3.3.3. KL Divergence between the True PDF and the Constructed PDF

The KL divergence is a non-symmetric measure of difference between two PDFs. For two PDFs  $p_1$  and  $p_0$ , it is defined as

$$D(p_1 \parallel p_0) = \int p_1(\mathbf{x}) \ln \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})} d\mathbf{x} \quad (48)$$

It is well known that  $D(p_1 \parallel p_0) \geq 0$  with equality if and only if  $p_1 = p_0$  almost everywhere [27]. As we mentioned in Section 3.3.2 the KL divergence determines the asymptotic performance for both detection and classification. Next we will show that under  $H_0$ ,  $p_T(\mathbf{t}; \hat{\boldsymbol{\theta}}) = p_T(\mathbf{t}; H_0)$  asymptotically, and similarly under  $H_1$ , within the family of PDFs in (38),  $p_T(\mathbf{t}; \hat{\boldsymbol{\theta}})$  is asymptotically the closest one to the true PDF in KL divergence. Similar results and arguments have been shown in [22], [36].

Assume that we observe IID  $\mathbf{T}^{(n)}$ 's with

$$\mathbf{T}^{(n)} = \begin{bmatrix} \mathbf{T}_1^{(n)} \\ \mathbf{T}_2^{(n)} \end{bmatrix} \quad (49)$$

for  $n=1, 2, \dots, L$ . Shortening the notation, we will write  $p_{\mathbf{T}^{(1)}, \mathbf{T}^{(2)}, \dots, \mathbf{T}^{(L)}}(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; \hat{\boldsymbol{\theta}})$  as  $p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; \hat{\boldsymbol{\theta}})$ . The constructed PDF can be easily extended to (see (10))

$$p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; \hat{\boldsymbol{\theta}}) = \exp \left[ \hat{\boldsymbol{\theta}}^T \sum_{n=1}^L \mathbf{t}^{(n)} - LK(\hat{\boldsymbol{\theta}}) + \ln p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; H_0) \right] \quad (50)$$

where the MLE  $\hat{\boldsymbol{\theta}}$  is obtained by solving

$$\frac{1}{L} \sum_{n=1}^L \mathbf{t}^{(n)} = \frac{\partial K(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \quad (51)$$

Now we consider two cases. First, for the true PDF under  $H_0$ , by the law of large numbers, it follows that

$$\frac{1}{L} \sum_{n=1}^L \mathbf{t}^{(n)} \xrightarrow{P} E_0(\mathbf{T}) \quad (52)$$

as  $L \rightarrow \infty$ . Note that

$$\frac{\partial K(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\mathbf{0}} = E_0(\mathbf{T}) \quad (53)$$

Since the solution of (51) is unique, asymptotically we have

$$\hat{\boldsymbol{\theta}} = \mathbf{0} \quad (54)$$

and hence,  $p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; \hat{\boldsymbol{\theta}}) = p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; H_0)$ .

Secondly, for the true PDF under  $H_1$ , by the law of large numbers, it follows that

$$\frac{1}{L} \sum_{n=1}^L \mathbf{t}^{(n)} \xrightarrow{P} E_1(\mathbf{T}) \quad (55)$$

as  $L \rightarrow \infty$ . Therefore, the MLE  $\hat{\boldsymbol{\theta}}$  asymptotically maximizes

$$\boldsymbol{\theta}^T E_1(\mathbf{T}) - K(\boldsymbol{\theta}) \quad (56)$$

We will denote the underlying true PDF under  $H_1$  as  $p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; H_1)$  and the PDF within the exponential family in (38) as  $p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; \boldsymbol{\theta})$ . Since from (18)

$$\ln \frac{p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; H_1)}{p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; \boldsymbol{\theta})} = - \left( \boldsymbol{\theta}^T \sum_{n=1}^L \mathbf{t}^{(n)} - LK(\boldsymbol{\theta}) \right) + \ln \frac{p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; H_1)}{p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; H_0)} \quad (57)$$

the KL divergence between the true PDF and the one in the exponential family is

$$\begin{aligned} & D(p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; H_1) \| p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; \boldsymbol{\theta})) \\ &= E_1 \left[ - \left( \boldsymbol{\theta}^T \sum_{n=1}^L \mathbf{t}^{(n)} - LK(\boldsymbol{\theta}) \right) + \ln \frac{p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; H_1)}{p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; H_0)} \right] \\ &= -L \left[ \boldsymbol{\theta}^T E_1(\mathbf{T}) - K(\boldsymbol{\theta}) \right] + D(p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; H_1) \| p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; H_0)) \end{aligned} \quad (58)$$

Since

$$D(p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; H_1) \| p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; H_0)) \quad (59)$$

is fixed,  $D(p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; H_1) \| p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; \boldsymbol{\theta}))$  is minimized by maximizing (56). This shows that given the exponential family in (38),  $p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; \hat{\boldsymbol{\theta}})$  is asymptotically the closest to  $p(\mathbf{t}^{(1)}, \mathbf{t}^{(2)}, \dots, \mathbf{t}^{(L)}; H_1)$  in KL divergence.

### 3.3.4. Examples – Distributed Detection

In this section, we compare our method with the clairvoyant GLRT for a specific detection problem. Since the true PDF does not necessarily belong to the exponential family in (8), for the clairvoyant GLRT, assume that we know the true PDF of  $\mathbf{T}$  under  $H_1$  belongs to a family of PDFs parameterized by some unknown parameters  $\boldsymbol{\alpha}$ . Note that the  $\boldsymbol{\theta}$  parameters in our method are constructed exponential PDFs, while  $\boldsymbol{\alpha}$  are parameters in the true PDF. Therefore, the clairvoyant GLRT provides an upper bound on GLRT performance, and it decides  $H_1$  if

$$\ln \frac{p_{\mathbf{T}}(\mathbf{t}; \hat{\boldsymbol{\alpha}})}{p_{\mathbf{T}}(\mathbf{t}; H_0)} > \tau \quad (60)$$

#### Example - Partially Observed Linear Model with Gaussian Noise

Suppose we have the linear model with

$$\mathbf{x} = \mathbf{H}\boldsymbol{\alpha} + \mathbf{w} \quad (61)$$

with

$$\begin{aligned} H_0 : \boldsymbol{\alpha} &= \mathbf{0} \\ H_1 : \boldsymbol{\alpha} &\neq \mathbf{0} \end{aligned} \quad (62)$$

where  $\mathbf{x}$  is an  $N \times 1$  vector of the underlying unobservable samples,  $\mathbf{H}$  is an  $N \times p$  observation matrix with full column rank,  $\boldsymbol{\alpha}$  is an  $p \times 1$  vector of the unknown signal amplitudes, and  $\mathbf{w}$  is an  $N \times 1$  vector of white Gaussian noise samples with known variance  $\sigma^2$ . We observe two sensor outputs

$$\mathbf{T}_1(\mathbf{x}) = \mathbf{H}_1^T \mathbf{x} \quad (63)$$

$$\mathbf{T}_2(\mathbf{x}) = \mathbf{H}_2^T \mathbf{x} \quad (64)$$

where  $\mathbf{H}_1$  is  $N \times q_1$  and  $\mathbf{H}_2$  is  $N \times q_2$ . Note that  $[\mathbf{H}_1 \mathbf{H}_2]$  does not have to be  $\mathbf{H}$ . This model is called the partially observed linear model.

Let  $\mathbf{G} = [\mathbf{H}_1 \ \mathbf{H}_2]$ . We assume that  $\mathbf{G}$  has full column rank so that there is no perfectly redundant measurements of the sensors. Then we have

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_1(\mathbf{x}) \\ \mathbf{T}_2(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \mathbf{H}_1^T \mathbf{x} \\ \mathbf{H}_2^T \mathbf{x} \end{bmatrix} = \mathbf{G}^T \mathbf{x} \quad (65)$$

Thus,  $\mathbf{T}$  is also Gaussian and

$$\mathbf{T} \sim N(0, \sigma^2 \mathbf{G}^T \mathbf{G}) \text{ under } H_0 \quad (66)$$

Let  $q = q_1 + q_2$ , and we can see that  $\mathbf{T}$  is  $q \times 1$ . As a result, we construct the PDF as in (38) with

$$K(\boldsymbol{\theta}) = \ln E_0 \left[ \exp(\boldsymbol{\theta}^T \mathbf{T}) \right] = \frac{1}{2} \sigma^2 \boldsymbol{\theta}^T \mathbf{G}^T \mathbf{G} \boldsymbol{\theta} \quad (67)$$

Hence the constructed PDF is

$$\begin{aligned} p_{\mathbf{T}}(\mathbf{t}; \boldsymbol{\theta}) &= \exp \left[ \boldsymbol{\theta}^T \mathbf{t} - K(\boldsymbol{\theta}) + \ln p_{\mathbf{T}}(\mathbf{t}; H_0) \right] \\ &= \frac{1}{(2\pi\sigma^2)^{\frac{q}{2}} \det^{\frac{1}{2}}(\mathbf{G}^T \mathbf{G})} \exp \left( -\frac{\mathbf{t}^T (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t}}{2\sigma^2} \right) \exp \left[ \boldsymbol{\theta}^T \mathbf{t} - \frac{1}{2} \sigma^2 \boldsymbol{\theta}^T \mathbf{G}^T \mathbf{G} \boldsymbol{\theta} \right] \end{aligned} \quad (68)$$

which can be simplified to

$$\mathbf{T} \sim N(\sigma^2 \mathbf{G}^T \mathbf{G} \boldsymbol{\theta}, \sigma^2 \mathbf{G}^T \mathbf{G}) \text{ under } H_1 \quad (69)$$

Note that  $\boldsymbol{\theta}$  is the vector of the unknown parameters in the constructed PDF, and it is different from the truly unknown parameters  $\boldsymbol{\alpha}$ . From (37) and (67), the MLE of  $\boldsymbol{\theta}$  satisfies

$$\mathbf{t} = \frac{\partial K(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \sigma^2 \mathbf{G}^T \mathbf{G} \boldsymbol{\theta} \quad (70)$$

So

$$\hat{\boldsymbol{\theta}} = \frac{1}{\sigma^2} (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t} \quad (71)$$

and the test statistic becomes

$$\hat{\boldsymbol{\theta}}^T \mathbf{t} - K(\hat{\boldsymbol{\theta}}) = \frac{1}{2\sigma^2} \mathbf{t}^T (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t} \quad (72)$$

Next we consider the clairvoyant GLRT. It is defined as the GLRT when the true PDF of  $\mathbf{T}$  under  $H_1$  is known except for the underlying unknown parameters  $\boldsymbol{\alpha}$ . It is considered as the suboptimal test by plugging the MLE of  $\boldsymbol{\alpha}$  into the true PDF parameterized by  $\boldsymbol{\alpha}$ . Since the constructed PDF may not be the true PDF, the clairvoyant GLRT requires more information than our method. From (65) we know that the true PDF is

$$\mathbf{T} \sim N(\mathbf{G}^T \mathbf{H}, \sigma^2 \mathbf{G}^T \mathbf{G}) \text{ under } H_1 \quad (73)$$

Note that (69) is the constructed PDF while (73) is the true PDF. For small signal, they are equivalent since their means are close to zero and they have the same covariance matrix. This verifies that our method works for the small signal case. We need to estimate  $\boldsymbol{\theta}$  in (69) or  $\boldsymbol{\alpha}$  in (73) to implement the PDF. We write the true PDF under  $H_1$  as  $p_T(\mathbf{t};)$ . The MLE of  $\boldsymbol{\alpha}$  is found by maximizing the true PDF given by (73)

$$\ln \frac{p_T(\mathbf{t}; \boldsymbol{\alpha})}{p_T(\mathbf{t}; \mathbf{0})} = -\frac{1}{2\sigma^2} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \boldsymbol{\alpha})^T (\mathbf{G}^T \mathbf{G})^{-1} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \boldsymbol{\alpha}) + \frac{1}{2\sigma^2} \mathbf{t}^T (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t} \quad (74)$$

If  $q \leq p$ , i.e., the length of  $\mathbf{t}$  is less than or equal to the length of  $\boldsymbol{\alpha}$ , then the MLE  $\hat{\boldsymbol{\alpha}}$  may not be unique. However, since  $(\mathbf{t} - \mathbf{G}^T \mathbf{H} \boldsymbol{\alpha})^T (\mathbf{G}^T \mathbf{G})^{-1} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \boldsymbol{\alpha}) \geq 0$ , we could always find  $\hat{\boldsymbol{\alpha}}$  such that  $\mathbf{t} = \mathbf{G}^T \mathbf{H} \hat{\boldsymbol{\alpha}}$  and hence  $(\mathbf{t} - \mathbf{G}^T \mathbf{H} \hat{\boldsymbol{\alpha}})^T (\mathbf{G}^T \mathbf{G})^{-1} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \hat{\boldsymbol{\alpha}}) = 0$ . Hence the clairvoyant GLRT statistic becomes

$$\ln \frac{p_T(\mathbf{t}; \hat{\boldsymbol{\alpha}})}{p_T(\mathbf{t}; H_0)} = \frac{1}{2\sigma^2} \mathbf{t}^T (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t} \quad (75)$$

which is the same as our test statistic (see (72)) when  $q \leq p$ .

If  $q > p$ , it can be shown that

$$\hat{\boldsymbol{\alpha}} = \left( \mathbf{H}^T \mathbf{G} (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{H} \right)^{-1} \mathbf{H}^T \mathbf{G} (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t} \quad (76)$$

and the clairvoyant GLRT statistic becomes

$$\ln \frac{p_T(\mathbf{t}; \hat{\boldsymbol{\alpha}})}{p_T(\mathbf{t}; H_0)} = \frac{\mathbf{t}^T (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{H} \left( \mathbf{H}^T \mathbf{G} (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{H} \right)^{-1} \mathbf{H}^T \mathbf{G} (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t}}{2\sigma^2} \quad (77)$$

while the constructed GLRT statistic is shown in (72).

### Example - Partially Observed Linear Model with Non-Gaussian Noise

The partially observed linear model remains the same as in the previous example except instead of assuming that  $\mathbf{w}$  is white Gaussian noise, we will assume that  $\mathbf{w}$  has a Gaussian mixture distribution with two components, i.e.,

$$\mathbf{w} \sim \pi N(0, \sigma_1^2 \mathbf{I}) + (1 - \pi) N(0, \sigma_2^2 \mathbf{I}) \quad (78)$$

where  $\pi$ ,  $\sigma_1^2$  and  $\sigma_2^2$  are known ( $0 < \pi < 1$ ). The following derivation can be easily extended when  $\mathbf{w} \sim \sum_{i=1}^L \pi_i N(0, \sigma_i^2 \mathbf{I})$ .

Since  $\mathbf{w}$  has a Gaussian mixture distribution,  $\mathbf{T} = \mathbf{G}^T \mathbf{x}$  is also Gaussian mixture distributed and

$$\mathbf{T} \sim \pi N(0, \sigma_1^2 \mathbf{G}^T \mathbf{G}) + (1 - \pi) N(0, \sigma_2^2 \mathbf{G}^T \mathbf{G}) \quad \text{under } H_0 \quad (79)$$

So we have

$$K(\boldsymbol{\theta}) = \ln E_0 \left[ \exp(\boldsymbol{\theta}^T \mathbf{t}) \right] = \ln \left( \pi e^{\frac{1}{2} \sigma_1^2 \mathbf{G}^T \mathbf{G}} + (1 - \pi) e^{\frac{1}{2} \sigma_2^2 \mathbf{G}^T \mathbf{G}} \right) \quad (80)$$

Hence the constructed PDF is

$$\begin{aligned} p_{\mathbf{T}}(\mathbf{t}; \boldsymbol{\theta}) &= \exp \left[ \boldsymbol{\theta}^T \mathbf{t} - K(\boldsymbol{\theta}) + \ln p_{\mathbf{T}}(\mathbf{t}; H_0) \right] \\ &= \left[ \frac{\pi}{(2\pi\sigma_1^2)^{\frac{q}{2}} \det^{\frac{1}{2}}(\mathbf{G}^T \mathbf{G})} \exp \left( -\frac{\mathbf{t}^T (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t}}{2\sigma_1^2} \right) + \frac{1 - \pi}{(2\pi\sigma_2^2)^{\frac{q}{2}} \det^{\frac{1}{2}}(\mathbf{G}^T \mathbf{G})} \exp \left( -\frac{\mathbf{t}^T (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t}}{2\sigma_2^2} \right) \right] \\ &\quad \cdot \exp(\boldsymbol{\theta}^T \mathbf{t}) / \left( \pi e^{\frac{1}{2} \sigma_1^2 \mathbf{G}^T \mathbf{G}} + (1 - \pi) e^{\frac{1}{2} \sigma_2^2 \mathbf{G}^T \mathbf{G}} \right) \end{aligned} \quad (81)$$

Although this constructed PDF cannot be further simplified, we can still find the MLE by solving

$$\mathbf{t} = \frac{\partial K(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \frac{\pi e^{\frac{1}{2} \sigma_1^2 \mathbf{G}^T \mathbf{G}} \cdot \sigma_1^2 \mathbf{G}^T \mathbf{G} + (1 - \pi) e^{\frac{1}{2} \sigma_2^2 \mathbf{G}^T \mathbf{G}} \cdot \sigma_2^2 \mathbf{G}^T \mathbf{G}}{\pi e^{\frac{1}{2} \sigma_1^2 \mathbf{G}^T \mathbf{G}} + (1 - \pi) e^{\frac{1}{2} \sigma_2^2 \mathbf{G}^T \mathbf{G}}} \quad (82)$$

Our test statistic is just

$$\hat{\boldsymbol{\theta}}^T \mathbf{t} - K(\hat{\boldsymbol{\theta}}) = \hat{\boldsymbol{\theta}}^T \mathbf{t} - \ln \left( \pi e^{\frac{1}{2} \sigma_1^2 \hat{\boldsymbol{\theta}}^T \mathbf{G}^T \mathbf{G} \hat{\boldsymbol{\theta}}} + (1 - \pi) e^{\frac{1}{2} \sigma_2^2 \hat{\boldsymbol{\theta}}^T \mathbf{G}^T \mathbf{G} \hat{\boldsymbol{\theta}}} \right) \quad (83)$$



where  $\hat{\boldsymbol{\theta}}$  satisfies (82). Although no analytical solution of the MLE of  $\boldsymbol{\theta}$  exists, it can be found using convex optimization techniques [30], [31]. Moreover, an analytical solution exists when  $\|\boldsymbol{\theta}\| \rightarrow 0$ . To see this, we will show that

$$\lim_{\|\boldsymbol{\theta}\| \rightarrow 0} \frac{\partial K(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} ./ \left( \pi \sigma_1^2 \mathbf{G}^T \mathbf{G} \boldsymbol{\theta} + (1 - \pi) \sigma_2^2 \mathbf{G}^T \mathbf{G} \boldsymbol{\theta} \right) = \mathbf{1} \quad (84)$$

where  $\frac{\partial K(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$  is shown in (82) and  $./$  means element-by-element division.

To prove (84), we have

$$\lim_{\|\boldsymbol{\theta}\| \rightarrow 0} \left( \pi e^{\frac{1}{2} \sigma_1^2 \boldsymbol{\theta}^T \mathbf{G}^T \mathbf{G} \boldsymbol{\theta}} + (1 - \pi) e^{\frac{1}{2} \sigma_2^2 \boldsymbol{\theta}^T \mathbf{G}^T \mathbf{G} \boldsymbol{\theta}} \right) = 1 \quad (85)$$

and

$$\lim_{\|\boldsymbol{\theta}\| \rightarrow 0} \left( \pi e^{\frac{1}{2} \sigma_1^2 \boldsymbol{\theta}^T \mathbf{G}^T \mathbf{G} \boldsymbol{\theta}} \cdot \sigma_1^2 \mathbf{G}^T \mathbf{G} \boldsymbol{\theta} + (1 - \pi) e^{\frac{1}{2} \sigma_2^2 \boldsymbol{\theta}^T \mathbf{G}^T \mathbf{G} \boldsymbol{\theta}} \cdot \sigma_2^2 \mathbf{G}^T \mathbf{G} \boldsymbol{\theta} \right) ./ \left( \pi \sigma_1^2 \mathbf{G}^T \mathbf{G} \boldsymbol{\theta} + (1 - \pi) \sigma_2^2 \mathbf{G}^T \mathbf{G} \boldsymbol{\theta} \right) = \mathbf{1} \quad (86)$$

by L'Hopital's rule. Dividing (86) by (85) and from (82), (84) is proved. As a result of (82) and (84), the MLE of  $\boldsymbol{\theta}$  satisfies

$$\mathbf{t} = \pi \sigma_1^2 \mathbf{G}^T \mathbf{G} + (1 - \pi) \sigma_2^2 \mathbf{G}^T \mathbf{G} \quad (87)$$

as  $\|\boldsymbol{\theta}\| \rightarrow 0$  and  $\hat{\boldsymbol{\theta}}$  can be easily found as

$$\hat{\boldsymbol{\theta}} = \frac{1}{\pi \sigma_1^2 + (1 - \pi) \sigma_2^2} (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t} \quad (88)$$

Since

$$\lim_{\|\boldsymbol{\theta}\| \rightarrow 0} K(\boldsymbol{\theta}) / \left( \frac{1}{2} \pi \sigma_1^2 \boldsymbol{\theta}^T \mathbf{G}^T \mathbf{G} \boldsymbol{\theta} + \frac{1}{2} (1 - \pi) \sigma_2^2 \boldsymbol{\theta}^T \mathbf{G}^T \mathbf{G} \boldsymbol{\theta} \right) = 1 \quad (89)$$

by using L'Hopital's rule twice, as  $\|\boldsymbol{\theta}\| \rightarrow 0$ , our test statistic becomes (see (83))

$$\hat{\boldsymbol{\theta}}^T \mathbf{t} - \left( \frac{1}{2} \pi \sigma_1^2 \hat{\boldsymbol{\theta}}^T \mathbf{G}^T \mathbf{G} \hat{\boldsymbol{\theta}} + \frac{1}{2} (1 - \pi) \sigma_2^2 \hat{\boldsymbol{\theta}}^T \mathbf{G}^T \mathbf{G} \hat{\boldsymbol{\theta}} \right) = \frac{1}{2(\pi \sigma_1^2 + (1 - \pi) \sigma_2^2)} \mathbf{t}^T (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t} \quad (90)$$

To find the clairvoyant GLRT statistic, we know that under  $H_1$  the true PDF is

$$p_T(\mathbf{t}; \mathbf{a}) = \frac{\pi}{(2\pi)^{q/2} \det^{1/2}(\sigma_1^2 \mathbf{G}^T \mathbf{G})} \exp \left[ -\frac{1}{2} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \mathbf{a})^T \frac{(\mathbf{G}^T \mathbf{G})^{-1}}{\sigma_1^2} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \mathbf{a}) \right] \\ + \frac{1-\pi}{(2\pi)^{q/2} \det^{1/2}(\sigma_2^2 \mathbf{G}^T \mathbf{G})} \exp \left[ -\frac{1}{2} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \mathbf{a})^T \frac{(\mathbf{G}^T \mathbf{G})^{-1}}{\sigma_2^2} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \mathbf{a}) \right] \quad (91)$$

Note the difference between (81) and (91) since (81) is the constructed PDF and (91) is the true PDF. The MLE of  $\mathbf{a}$  is found by maximizing (91) over  $\mathbf{a}$ .

When  $q \leq p$ , the MLE of  $\mathbf{a}$  may not be unique but satisfies  $\mathbf{t} = \mathbf{G}^T \mathbf{H} \hat{\mathbf{a}}$ . As a result,  $p_T(\mathbf{t}; \hat{\mathbf{a}})$  is a constant and the clairvoyant GLRT statistic becomes

$$-\ln p_T(\mathbf{t}; \mathbf{0}) \quad (92)$$

Note that since  $p_T(\mathbf{t}; \mathbf{0})$  is decreasing as  $\mathbf{t}^T (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t}$  increases, the clairvoyant GLRT statistic becomes

$$\mathbf{t}^T (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t} \quad (93)$$

which is the same as our test statistic (with only a positive scale factor) as  $\|\mathbf{0}\| \rightarrow 0$  (see (72)). However, for large signal, our test statistic in (83) is not equivalent to the clairvoyant GLRT statistic in (93). This example shows that our method may not offer the suboptimal performance as the clairvoyant GLRT does for large signal.

When  $q > p$ , it can be shown that

$$\hat{\mathbf{a}} = \left( \mathbf{H}^T \mathbf{G} (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{H} \right)^{-1} \mathbf{H}^T \mathbf{G} (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t} \quad (94)$$

and the clairvoyant GLRT statistic becomes

$$\frac{\pi}{(\sigma_1^2)^{q/2}} \exp \left[ -\frac{1}{2} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \hat{\mathbf{a}})^T \frac{(\mathbf{G}^T \mathbf{G})^{-1}}{\sigma_1^2} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \hat{\mathbf{a}}) \right] \\ + \frac{1-\pi}{(\sigma_2^2)^{q/2}} \exp \left[ -\frac{1}{2} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \hat{\mathbf{a}})^T \frac{(\mathbf{G}^T \mathbf{G})^{-1}}{\sigma_2^2} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \hat{\mathbf{a}}) \right] \quad (95)$$

Note that the noise in (78) is uncorrelated but not independent. We next consider a general case when the noise can be correlated with a Gaussian mixture PDF

$$\mathbf{w} \sim \pi N(0, \mathbf{C}_1) + (1 - \pi) N(0, \mathbf{C}_2) \quad (96)$$

It can be shown that similar to (83), our test statistic is

$$\hat{\boldsymbol{\theta}}^T \mathbf{t} - \ln \left( \pi e^{\frac{1}{2} \hat{\boldsymbol{\theta}}^T \mathbf{G}^T \mathbf{C}_1 \mathbf{G} \hat{\boldsymbol{\theta}}} + (1 - \pi) e^{\frac{1}{2} \hat{\boldsymbol{\theta}}^T \mathbf{G}^T \mathbf{C}_2 \mathbf{G} \hat{\boldsymbol{\theta}}} \right) \quad (97)$$

and the clairvoyant GLRT statistic is

$$-\ln \left( \frac{\pi}{\det^{1/2}(\mathbf{C}_1)} \exp \left[ -\frac{1}{2} \mathbf{t}^T (\mathbf{G}^T \mathbf{C}_1 \mathbf{G})^{-1} \mathbf{t} \right] + \frac{1 - \pi}{\det^{1/2}(\mathbf{C}_2)} \exp \left[ -\frac{1}{2} \mathbf{t}^T (\mathbf{G}^T \mathbf{C}_2 \mathbf{G})^{-1} \mathbf{t} \right] \right) \quad (98)$$

when  $q \leq p$ .

When  $q > p$ , the MLE of  $\boldsymbol{\alpha}$  is not in closed form, and hence we write the clairvoyant GLRT statistic as

$$\begin{aligned} \max_{\boldsymbol{\alpha}} \left[ \frac{\pi}{\det^{1/2}(\mathbf{G}^T \mathbf{C}_1 \mathbf{G})} \exp \left[ -\frac{1}{2} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \boldsymbol{\alpha})^T (\mathbf{G}^T \mathbf{C}_1 \mathbf{G})^{-1} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \boldsymbol{\alpha}) \right] \right. \\ \left. + \frac{1 - \pi}{\det^{1/2}(\mathbf{G}^T \mathbf{C}_2 \mathbf{G})} \exp \left[ -\frac{1}{2} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \boldsymbol{\alpha})^T (\mathbf{G}^T \mathbf{C}_2 \mathbf{G})^{-1} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \boldsymbol{\alpha}) \right] \right] \quad (99) \end{aligned}$$

### Distributed Detection Summary

We have considered the partially observed linear model with both Gaussian and non-Gaussian noise. Table 1 compares our test statistic with the clairvoyant GLRT.

- 1) In Gaussian noise,  $\mathbf{w} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$ . The test statistics are exactly the same for  $q \leq p$ .
- 2) In uncorrelated non-Gaussian noise,  $\mathbf{w} \sim \pi N(\mathbf{0}, \sigma_1^2 \mathbf{I}) + (1 - \pi) N(\mathbf{0}, \sigma_2^2 \mathbf{I})$ . The test statistics are the same as  $\boldsymbol{\theta} \rightarrow \mathbf{0}$  for  $q \leq p$ .
- 3) In correlated non-Gaussian noise,  $\mathbf{w} \sim \pi N(\mathbf{0}, \mathbf{C}_1) + (1 - \pi) N(\mathbf{0}, \mathbf{C}_2)$ . Although we cannot show the equivalence between these two test statistics, we will see in later that their performance appears to be the same.

**Table 1. Comparison of Our Test Statistic and the Clairvoyant GLRT**

	Our Method	Clairvoyant GLRT ( $q \leq p$ )	Clairvoyant GLRT ( $q > p$ )
Gaussian Noise	$\mathbf{t}^T (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t}$	$\mathbf{t}^T (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t}$	$\mathbf{t}^T (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{H} (\mathbf{H}^T \mathbf{G} (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{G} (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t}$
Uncorrelated Non-Gaussian Noise	$\max_{\theta} \left[ \theta^T \mathbf{t} - \ln \left( \pi e^{\frac{1}{2} \sigma_1^2 \theta^T \mathbf{G}^T \mathbf{G} \theta} + (1 - \pi) e^{\frac{1}{2} \sigma_2^2 \theta^T \mathbf{G}^T \mathbf{G} \theta} \right) \right]$	$\mathbf{t}^T (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t}$	$\frac{\pi}{(\sigma_1^2)^{q/2}} \exp \left[ -\frac{1}{2} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \alpha)^T \frac{(\mathbf{G}^T \mathbf{G})^{-1}}{\sigma_1^2} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \alpha) \right] + \frac{1 - \pi}{(\sigma_2^2)^{q/2}} \exp \left[ -\frac{1}{2} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \alpha)^T \frac{(\mathbf{G}^T \mathbf{G})^{-1}}{\sigma_2^2} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \alpha) \right]$
Correlated Non-Gaussian Noise	$\max_{\theta} \left[ \theta^T \mathbf{t} - \ln \left( \pi e^{\frac{1}{2} \theta^T \mathbf{G}^T \mathbf{C}_1 \mathbf{G} \theta} + (1 - \pi) e^{\frac{1}{2} \theta^T \mathbf{G}^T \mathbf{C}_2 \mathbf{G} \theta} \right) \right]$	$-\ln \left( \frac{\pi}{\det^{1/2}(\mathbf{C}_1)} \exp \left[ -\frac{1}{2} \mathbf{t}^T (\mathbf{G}^T \mathbf{C}_1 \mathbf{G})^{-1} \mathbf{t} \right] + \frac{1 - \pi}{\det^{1/2}(\mathbf{C}_2)} \exp \left[ -\frac{1}{2} \mathbf{t}^T (\mathbf{G}^T \mathbf{C}_2 \mathbf{G})^{-1} \mathbf{t} \right] \right)$	$\max_{\alpha} \left[ \frac{\pi}{\det^{1/2}(\mathbf{G}^T \mathbf{C}_1 \mathbf{G})} \exp \left[ -\frac{1}{2} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \alpha)^T (\mathbf{G}^T \mathbf{C}_1 \mathbf{G})^{-1} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \alpha) \right] + \frac{1 - \pi}{\det^{1/2}(\mathbf{G}^T \mathbf{C}_2 \mathbf{G})} \exp \left[ -\frac{1}{2} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \alpha)^T (\mathbf{G}^T \mathbf{C}_2 \mathbf{G})^{-1} (\mathbf{t} - \mathbf{G}^T \mathbf{H} \alpha) \right] \right]$

### 3.3.5. Examples-Distributed Classification

In this section, we compare our method with the estimated MAP classifier for some classification problems. The estimated MAP classifier assumes that the PDF of  $\mathbf{T}$  under  $H_i$  is known except for some unknown underlying parameters  $\mathbf{a}_i$ . We assume equal prior probability of the candidate hypothesis, i.e.,  $p(H_1) = \dots = p(H_M) = 1/M$ . So the estimated MAP classifier reduces to the estimated maximum likelihood classifier [1], which finds the MLE of  $\mathbf{a}_i$  and chooses  $H_i$  for which the following is maximum over  $i$ :

$$p_T(\mathbf{t}; \hat{\mathbf{a}}_i) \quad (100)$$

where  $\hat{\mathbf{a}}_i$  is the MLE of  $\mathbf{a}_i$ .

#### Example: Linear Model with Known Variance

Consider the following classification model:

$$H_i: \mathbf{x} = A_i \mathbf{s}_i + \mathbf{w} \quad (101)$$

where  $\mathbf{s}_i$  is an  $N \times 1$  known signal vector with the same length as  $\mathbf{x}$  and depends upon the class,  $A_i$  is the unknown signal amplitude, and  $\mathbf{w}$  is white Gaussian noise with known variance  $\sigma^2$ . Assume that instead of observing  $\mathbf{x}$ , we can only observe the measurements of two sensors

$$\mathbf{T}_1 = \mathbf{H}_1^T \mathbf{x} \quad (102)$$

$$\mathbf{T}_2 = \mathbf{H}_2^T \mathbf{x} \quad (103)$$

where  $\mathbf{H}_1$  is  $N \times q_1$  and  $\mathbf{H}_2$  is  $N \times q_2$ . Here  $q_1$  and  $q_2$  are the length for vectors  $\mathbf{T}_1$  and  $\mathbf{T}_2$  respectively. We can write (103) as

$$\mathbf{T} = \mathbf{G}^T \mathbf{x} \quad (104)$$

by letting

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_1 \\ \mathbf{T}_2 \end{bmatrix} \quad (105)$$

and

$$\mathbf{G} = [\mathbf{H}_1 \ \mathbf{H}_2] \quad (106)$$

where  $\mathbf{G}$  is  $N \times (q_1 + q_2)$  with  $q_1 + q_2 \leq N$ . We assume that  $\mathbf{G}$  has full column rank so that there are no perfectly redundant measurements of the sensors. Note that  $\mathbf{G}$  can be any matrix with full column rank.

Let  $H_0$  be the reference hypothesis when there is noise only, i.e.,

$$H_0: \mathbf{x} = \mathbf{w} \quad (107)$$

Since  $\mathbf{x}$  is Gaussian under  $H_0$ , according to (104),  $\mathbf{T}$  is also Gaussian and

$$\mathbf{T} \sim N(\mathbf{0}, \sigma^2 \mathbf{G}^T \mathbf{G}) \quad (108)$$

under  $H_0$ . We construct the PDF under  $H_i$  as in (13) with

$$K(\boldsymbol{\theta}_i) = \ln E_0 \left[ \exp(\boldsymbol{\theta}_i^T \mathbf{T}) \right] = \frac{1}{2} \sigma^2 \boldsymbol{\theta}_i^T \mathbf{G}^T \mathbf{G} \boldsymbol{\theta}_i \quad (109)$$

Hence the constructed PDF is

$$\begin{aligned} p_{\mathbf{T}}(\mathbf{t}; \boldsymbol{\theta}_i) &= \exp \left[ \boldsymbol{\theta}_i^T \mathbf{t} - K(\boldsymbol{\theta}_i) + \ln p_{\mathbf{T}}(\mathbf{t}; H_0) \right] \\ &= \frac{1}{\left( 2\pi\sigma^2 \right)^{\frac{q_1+q_2}{2}} \det^{\frac{1}{2}}(\mathbf{G}^T \mathbf{G})} \exp \left( -\frac{\mathbf{t}^T (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t}}{2\sigma^2} \right) \exp \left[ \boldsymbol{\theta}_i^T \mathbf{t} - \frac{1}{2} \sigma^2 \boldsymbol{\theta}_i^T \mathbf{G}^T \mathbf{G} \boldsymbol{\theta}_i \right] \end{aligned} \quad (110)$$

which can be simplified as

$$\mathbf{T} \sim N(\sigma^2 \mathbf{G}^T \mathbf{G}_i, \sigma^2 \mathbf{G}^T \mathbf{G}) \text{ under } H_i \quad (111)$$

The next step is to find the MLE of  $\boldsymbol{\theta}_i$ . Note that the MLE of  $\boldsymbol{\theta}_i$  is found by maximizing  $\boldsymbol{\theta}_i^T \mathbf{t} - K(\boldsymbol{\theta}_i)$  over  $\boldsymbol{\theta}_i$ . If this optimization procedure is carried out without any constraints, then  $\hat{\boldsymbol{\theta}}_i$  would be the same for all  $i$ . Hence we need some implicit constraints in finding the MLE. Since  $\boldsymbol{\theta}_i$  represents the signal under  $H_i$ , we should have

$$\boldsymbol{\theta}_i = A_i \mathbf{G}^T \mathbf{s}_i = E_{H_i}(\mathbf{T}) \quad (112)$$

which is the mean of  $\mathbf{T}$  under  $H_i$ . As a result, (111) can be written as

$$\mathbf{T} \sim N(\sigma^2 A_i \mathbf{G}^T \mathbf{G} \mathbf{G}^T \mathbf{s}_i, \sigma^2 \mathbf{G}^T \mathbf{G}) \text{ under } H_i \quad (113)$$

Thus, instead of finding the MLE of  $\boldsymbol{\theta}_i$  by maximizing

$$\boldsymbol{\theta}_i^T \mathbf{t} - K(\boldsymbol{\theta}_i) = \boldsymbol{\theta}_i^T \mathbf{t} - \frac{1}{2} \sigma^2 \boldsymbol{\theta}_i^T \mathbf{G}^T \mathbf{G} \boldsymbol{\theta}_i \quad (114)$$

with the constraint in (112), we can find the MLE of  $A_i$  in (113) (since  $\mathbf{s}_i$  is assumed known) and then plug it into (112). It can be shown that

$$\hat{A}_i = \frac{\mathbf{s}_i^T \mathbf{G} \mathbf{t}}{\sigma^2 \mathbf{s}_i^T \mathbf{G} \mathbf{G}^T \mathbf{G} \mathbf{G}^T \mathbf{s}_i} \quad (115)$$

and

$$\hat{\boldsymbol{\theta}}_i = \frac{\mathbf{G}^T \mathbf{s}_i \mathbf{s}_i^T \mathbf{G} \mathbf{t}}{\sigma^2 \mathbf{s}_i^T \mathbf{G} \mathbf{G}^T \mathbf{G} \mathbf{G}^T \mathbf{s}_i} \quad (116)$$

Hence by removing the constant factors, the test statistic of our classifier for  $H_i$  is

$$\frac{(\mathbf{s}_i^T \mathbf{G} \mathbf{t})^2}{(\mathbf{G}^T \mathbf{s}_i)^T \mathbf{G}^T \mathbf{G} (\mathbf{G}^T \mathbf{s}_i)} \quad (117)$$

according to (47).

Next we consider the estimate MAP classifier. In this case, we assume that we know the true PDF except for  $A_i$

$$\mathbf{T} \sim N(A_i \mathbf{G}^T \mathbf{s}_i, \sigma^2 \mathbf{G}^T \mathbf{G}) \text{ under } H_i \quad (118)$$

Note that (118) is the true PDF of  $\mathbf{T}$  under  $H_i$  and (113) is the constructed PDF. It can be shown that the MLE of  $A_i$  in the true PDF under  $H_i$  is

$$\hat{A}_i = \frac{\mathbf{s}_i^T \mathbf{G} (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t}}{\mathbf{s}_i^T \mathbf{G} (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{s}_i} \quad (119)$$

By removing the constant terms, the test statistic of the estimated MAP classifier for  $H_i$  is

$$\frac{(\mathbf{s}_i^T \mathbf{G} (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t})^2}{(\mathbf{G}^T \mathbf{s}_i)^T (\mathbf{G}^T \mathbf{G})^{-1} (\mathbf{G}^T \mathbf{s}_i)} \quad (120)$$

according to (100). Note that (115) and (119) are different because (115) is the MLE of  $A_i$  under the constructed PDF and (119) is the MLE of  $A_i$  under the true PDF. Also note that if  $\mathbf{G}^T \mathbf{G}$  is a scaled identity matrix, test statistics in (117) and (120) are equivalent, and hence our method coincides with the estimated MAP classifier.

#### Example: Linear Model with Unknown Variance

To extend the above example, we consider the above linear model with unknown noise variance  $\sigma^2$ . As we have shown in (113), the constructed PDF is still

$$\mathbf{T} \sim N(\sigma^2 A_i \mathbf{G}^T \mathbf{G} \mathbf{G}^T \mathbf{s}_i, \sigma^2 \mathbf{G}^T \mathbf{G}) \text{ under } H_i \quad (121)$$

except for that  $\sigma^2$  is unknown. Let  $B_i = \sigma^2 A_i$ , we have

$$\mathbf{T} \sim N(B_i \mathbf{G}^T \mathbf{G} \mathbf{G}^T \mathbf{s}_i, \sigma^2 \mathbf{G}^T \mathbf{G}) \text{ under } H_i \quad (122)$$

Instead of finding the MLEs of  $A_i$  and  $\sigma^2$ , we can equivalently find the MLEs of  $B_i$  and  $\sigma^2$ . Let  $\mathbf{h}_i = \mathbf{G}^T \mathbf{G} \mathbf{G}^T \mathbf{s}_i$  and  $\mathbf{C} = \mathbf{G}^T \mathbf{G}$ . It can be shown that

$$\hat{B}_i = (\mathbf{h}_i^T \mathbf{C}^{-1} \mathbf{h}_i)^{-1} \mathbf{h}_i^T \mathbf{C}^{-1} \mathbf{t} \quad (123)$$

and

$$\hat{\sigma}^2 = \frac{1}{p_1 + p_2} (\mathbf{t} - \mathbf{h}_i \hat{B}_i)^T \mathbf{C}^{-1} (\mathbf{t} - \mathbf{h}_i \hat{B}_i) \quad (124)$$

By removing the constant factors, it can also be shown that the test statistic is equivalent to

$$\frac{\mathbf{t}^T \mathbf{C}^{-1} \mathbf{h}_i (\mathbf{h}_i^T \mathbf{C}^{-1} \mathbf{h}_i)^{-1} \mathbf{h}_i^T \mathbf{C}^{-1} \mathbf{t}}{\mathbf{t}^T [\mathbf{C}^{-1} - \mathbf{C}^{-1} \mathbf{h}_i (\mathbf{h}_i^T \mathbf{C}^{-1} \mathbf{h}_i)^{-1} \mathbf{h}_i^T \mathbf{C}^{-1}] \mathbf{t}} \quad (125)$$

Next we consider the estimated MAP classifier. The true PDF is still

$$\mathbf{T} \sim N(\mathbf{A}_i \mathbf{G}^T \mathbf{s}_i, \sigma^2 \mathbf{G}^T \mathbf{G}) \quad \text{under } H_i \quad (126)$$

but with unknown  $\mathbf{A}_i$  and  $\sigma^2$ . Let  $\mathbf{g}_i = \mathbf{G}^T \mathbf{s}_i$  and  $\mathbf{C} = \mathbf{G}^T \mathbf{G}$ . Similar to (123), (124) and (125), it can be shown that

$$\hat{\mathbf{A}}_i = (\mathbf{g}_i^T \mathbf{C}^{-1} \mathbf{g}_i)^{-1} \mathbf{g}_i^T \mathbf{C}^{-1} \mathbf{t} \quad (127)$$

$$\hat{\sigma}^2 = \frac{1}{p_1 + p_2} (\mathbf{t} - \mathbf{g}_i \hat{\mathbf{A}}_i)^T \mathbf{C}^{-1} (\mathbf{t} - \mathbf{g}_i \hat{\mathbf{A}}_i) \quad (128)$$

and the test statistic of the estimated MAP classifier is

$$\frac{\mathbf{t}^T \mathbf{C}^{-1} \mathbf{g}_i (\mathbf{g}_i^T \mathbf{C}^{-1} \mathbf{g}_i)^{-1} \mathbf{h}_i^T \mathbf{C}^{-1} \mathbf{t}}{\mathbf{t}^T [\mathbf{C}^{-1} - \mathbf{C}^{-1} \mathbf{g}_i (\mathbf{g}_i^T \mathbf{C}^{-1} \mathbf{g}_i)^{-1} \mathbf{g}_i^T \mathbf{C}^{-1}] \mathbf{t}} \quad (129)$$

Note that if  $\mathbf{G}^T \mathbf{G}$  is a scaled identity matrix, since  $\mathbf{h}_i = \mathbf{G}^T \mathbf{G} \mathbf{g}_i$ , the test statistics in (125) and (129) are equivalent. Hence our method is exactly the same as the estimated MAP classifier if  $\mathbf{G}^T \mathbf{G}$  is a scaled identity matrix.

### Distributed Classification Summary

We have considered a linear model both known and unknown noise variance. Table 2 compares our test statistic with the estimated MAP classifier. If  $\mathbf{G}^T \mathbf{G}$  is a scaled identity matrix, our method and the estimated MAP classifier are identical. Note that this is the case when all the columns in  $\mathbf{G}$  are orthogonal and have same power, such as the demodulation of M-ary orthogonal signals in communication theory.

**Table 2. Comparison of Our Test Statistic and the Estimated MAP Classifier**

	Our Method	Estimated MAP
Known $\sigma^2$	$\frac{(\mathbf{s}_i^T \mathbf{G} \mathbf{t})^2}{(\mathbf{G}^T \mathbf{s}_i)^T \mathbf{G}^T \mathbf{G} (\mathbf{G}^T \mathbf{s}_i)}$	$\frac{(\mathbf{s}_i^T \mathbf{G} (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{t})^2}{(\mathbf{G}^T \mathbf{s}_i)^T (\mathbf{G}^T \mathbf{G})^{-1} (\mathbf{G}^T \mathbf{s}_i)}$
Unknown $\sigma^2$	$\frac{\mathbf{t}^T \mathbf{C}^{-1} \mathbf{h}_i (\mathbf{h}_i^T \mathbf{C}^{-1} \mathbf{h}_i)^{-1} \mathbf{h}_i^T \mathbf{C}^{-1} \mathbf{t}}{\mathbf{t}^T [\mathbf{C}^{-1} - \mathbf{C}^{-1} \mathbf{h}_i (\mathbf{h}_i^T \mathbf{C}^{-1} \mathbf{h}_i)^{-1} \mathbf{h}_i^T \mathbf{C}^{-1}] \mathbf{t}}$	$\frac{\mathbf{t}^T \mathbf{C}^{-1} \mathbf{g}_i (\mathbf{g}_i^T \mathbf{C}^{-1} \mathbf{g}_i)^{-1} \mathbf{h}_i^T \mathbf{C}^{-1} \mathbf{t}}{\mathbf{t}^T [\mathbf{C}^{-1} - \mathbf{C}^{-1} \mathbf{g}_i (\mathbf{g}_i^T \mathbf{C}^{-1} \mathbf{g}_i)^{-1} \mathbf{g}_i^T \mathbf{C}^{-1}] \mathbf{t}}$

where  $\mathbf{h}_i = \mathbf{G}^T \mathbf{G} \mathbf{g}_i$ ,  $\mathbf{g}_i = \mathbf{G}^T \mathbf{s}_i$  and  $\mathbf{C} = \mathbf{G}^T \mathbf{G}$ .



### 3.3.6. Simulations

#### Simulation: Distributed Detection

Since our test statistic coincides with the clairvoyant GLRT under Gaussian noise for  $q \leq p$  as shown in subsection 3.3.4, we will only compare the performances under non-Gaussian noise (both uncorrelated noise as in (78) and correlated noise as in (96)). Consider the model where

$$x[n] = A_1 + A_2 r^n + A_3 \cos(2\pi f_0 n + \phi) + w[n] \quad (130)$$

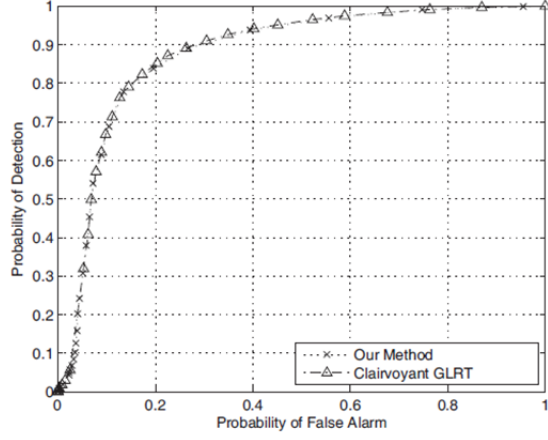
for  $n = 0, 1, \dots, N-1$  with known damping factor  $r \in (0, 1)$  and frequency  $f_0$  but unknown amplitudes  $A_1, A_2, A_3$  and phase  $\phi$ . This is a linear model as in (61) where

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & r & \cos(2\pi f_0) & \sin(2\pi f_0) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & r^{N-1} & \cos(2\pi f_0(N-1)) & \sin(2\pi f_0(N-1)) \end{bmatrix} \quad (131)$$

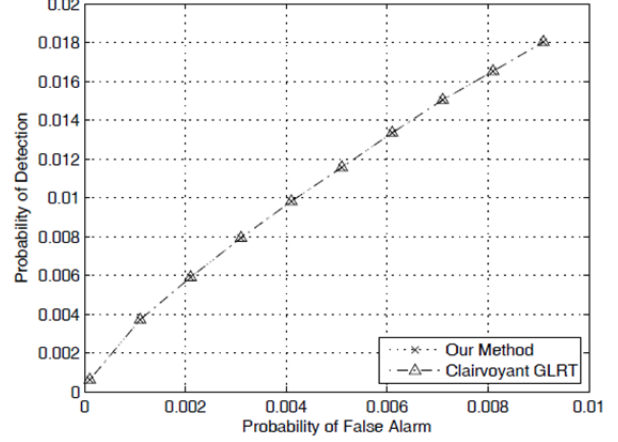
and  $\mathbf{a} = [A_1 \ A_2 \ A_3 \cos \phi \ -A_3 \sin \phi]^T$ .

Let  $\mathbf{w}$  have an uncorrelated Gaussian mixture distribution as in (78). For the partially observed linear model, we observe two sensor outputs as in (63, 64). We compare the GLRT in (83) with the clairvoyant GLRT in (93). Note that the MLE of  $\mathbf{\theta}$  in (83) is found numerically, not by the asymptotic approximation in (88).

In the simulation, we use  $N = 20$ ,  $A_1 = 2$ ,  $A_2 = 3$ ,  $A_3 = 4$ ,  $\phi = \pi/4$ ,  $r = 0.95$ ,  $f_0 = 0.34$ ,  $\pi = 0.9$ ,  $\sigma_1^2 = 50$ ,  $\sigma_2^2 = 500$ , and  $\mathbf{H}_1$  and  $\mathbf{H}_2$  are the first and third columns in  $\mathbf{H}$  respectively, i.e.,  $\mathbf{H}_1 = [1 \ 1 \ \dots \ 1]^T$ ,  $\mathbf{H}_2 = [1 \ \cos(2\pi f_0) \ \dots \ \cos(2\pi f_0(N-1))]^T$ . Hence, only the DC level is sensed by one sensor and the in-phase component of the sinusoid is sensed by the other sensor. As shown in Figure 3, the performances are almost the same (even for very small false alarm rate) which justifies their equivalence when  $\|\mathbf{\theta}\| \rightarrow 0$  as shown in Section 3.2.4.



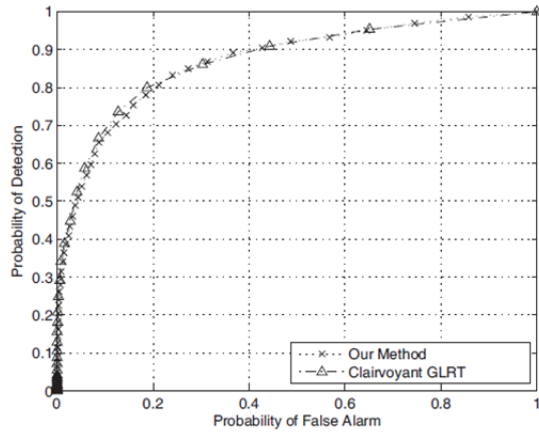
(a) False alarm rate between 0 and 1



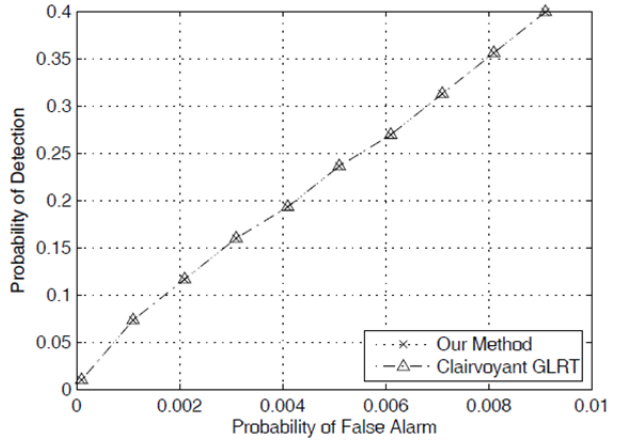
(b) False alarm rate between 0.0001 and 0.01

**Figure 3. GLRT ROC Curves for Unrelated Gaussian Mixture Noise Constructed PDF and the Clairvoyant GLRT**

Next for the same model in (130), let  $\mathbf{w}$  have a correlated Gaussian mixture distribution as in (96). We compare performances of the GLRT using the constructed PDF as in (97) and the clairvoyant GLRT as in (98). We use  $N = 20$ ,  $A_1 = 3$ ,  $A_2 = 4$ ,  $A_3 = 3$ ,  $\phi = \pi/7$ ,  $r = 0.9$ ,  $f_0 = 0.46$ ,  $\pi = 0.7$ ,  $\mathbf{H}_1 = [1 \ 1 \ \dots \ 1]^T$ ,  $\mathbf{H}_2 = [1 \ \cos(2\pi f_0) \ \dots \ \cos(2\pi f_0(N-1))]^T$ . The covariance matrices  $\mathbf{C}_1$ ,  $\mathbf{C}_2$  are generated using  $\mathbf{C}_1 = \mathbf{R}_1^T \mathbf{R}_1$ ,  $\mathbf{C}_2 = \mathbf{R}_2^T \mathbf{R}_2$ , where  $\mathbf{R}_1$ ,  $\mathbf{R}_2$  are full rank  $N \times N$  matrices. As shown in Figure 4, the performances are still very similar even for small false alarm rate.



(a) False alarm rate between 0 and 1



(b) False alarm rate between 0.0001 and 0.01

**Figure 4. GLRT ROC Curves for Correlated Gaussian Mixture Noise Constructed PDF and the Clairvoyant GLRT**

### Simulation: Distributed Classification

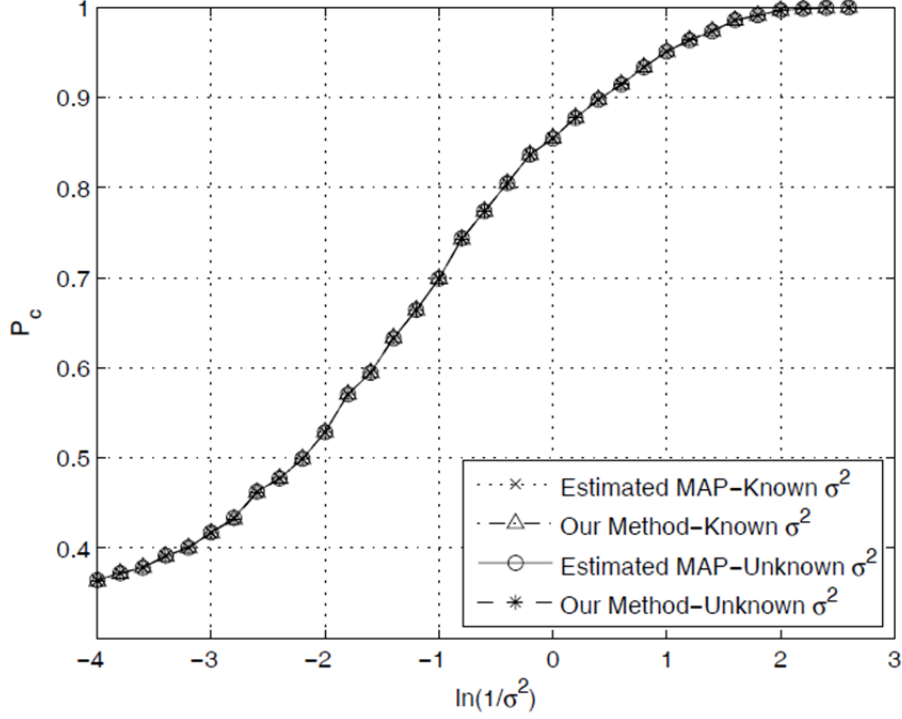
For the model in (101) we first consider a case when  $\mathbf{G}^T \mathbf{G}$  is approximately a scaled identity matrix. Let  $A_1 = 0.4$ ,  $A_2 = 1.2$ ,  $A_3 = 0.9$  and

$$\begin{aligned} s_1(n) &= \cos(2\pi f_1 n) \\ s_2(n) &= \cos(2\pi f_2 n) \\ s_3(n) &= \cos(2\pi f_3 n) \end{aligned} \tag{132}$$

where  $n = 0, 1, \dots, N-1$  with  $N = 25$ , and  $f_1 = 0.14$ ,  $f_2 = 0.34$ ,  $f_3 = 0.41$ . Let  $p(H_1) = p(H_2) = p(H_3) = 1/3$ . Assume that there are two sensors, each with an observation matrix as follows respectively:

$$\begin{aligned} \mathbf{H}_1 &= \begin{bmatrix} 1 & \cos(2\pi f_1) & \cdots & \cos(2\pi f_1(N-1)) \\ 1 & \cos(2\pi f_2) & \cdots & \cos(2\pi f_2(N-1)) \end{bmatrix}^T \\ \mathbf{H}_2 &= \begin{bmatrix} 1 & \cos(2\pi f_3) & \cdots & \cos(2\pi f_3(N-1)) \end{bmatrix}^T \end{aligned} \tag{133}$$

We use (117) and (120) as our test statistics for the two methods respectively, when  $\sigma^2$  is known. Test statistics in (125) and (129) are used when  $\sigma^2$  is unknown. The probabilities of correct classification are plotted versus  $\ln(1/\sigma^2)$  in Figure 5. We see that our method has the same performance with the estimated MAP classifier with known or unknown  $\sigma^2$ , and probabilities of correct classification goes to 1 as  $\sigma^2 \rightarrow 0$ .



**Figure 5. Probability of Correct Classification for Both Methods**

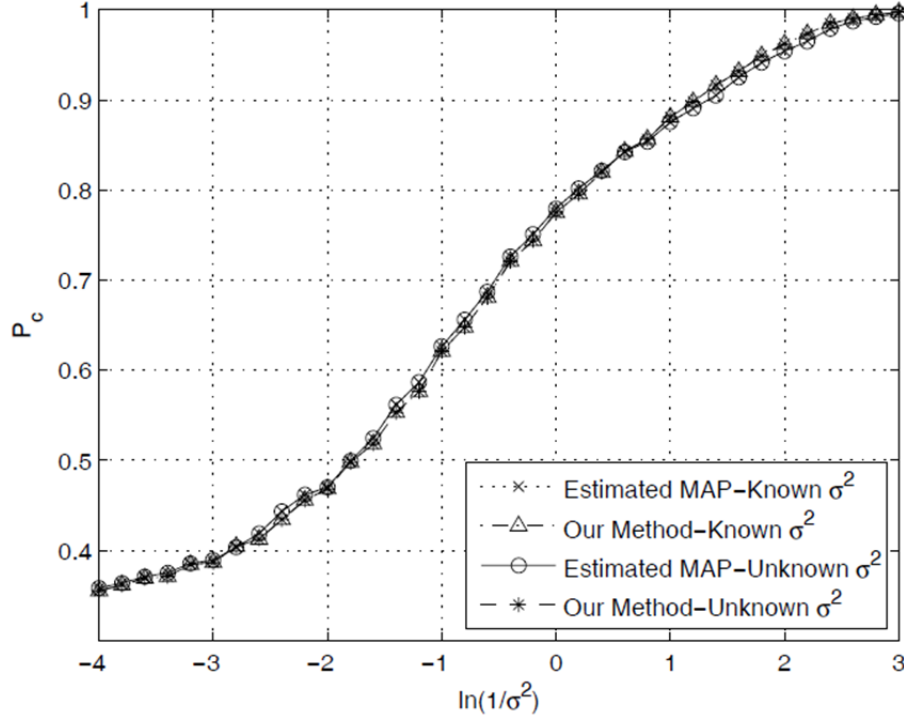
Next we consider a case when  $\mathbf{G}^T \mathbf{G}$  is not a scaled identity matrix. Let  $A_1 = 0.5$ ,  $A_2 = 1$ ,  $A_3 = 1$  and

$$\begin{aligned} s_1(n) &= \cos(2\pi f_1 n) + 1 \\ s_2(n) &= \cos(2\pi f_2 n) + 0.5 \\ s_3(n) &= \cos(2\pi f_3 n) \end{aligned} \quad (134)$$

where  $n = 0, 1, \dots, N-1$  with  $N = 20$ , and  $f_1 = 0.17$ ,  $f_2 = 0.28$ ,  $f_3 = 0.45$ . Let  $p(H_1) = p(H_2) = p(H_3) = 1/3$ . Assume that there are three sensors (this is an extension of the two sensor assumption), each with an observation matrix as follows respectively:

$$\begin{aligned} \mathbf{H}_1 &= [1 \quad 1 \quad \dots \quad 1]^T \\ \mathbf{H}_2 &= \begin{bmatrix} 1 & \cos(2\pi f_1) & \dots & \cos(2\pi f_1(N-1)) \\ 1 & \cos(2\pi f_2) & \dots & \cos(2\pi f_2(N-1)) \end{bmatrix}^T \\ \mathbf{H}_3 &= \begin{bmatrix} 1 & \cos(2\pi(f_3 + 0.02)) & \dots & \cos(2\pi(f_3 + 0.02)(N-1)) \end{bmatrix}^T \end{aligned} \quad (135)$$

Note that in  $\mathbf{H}_3$ , we set the frequency to  $f_3 + 0.02$ . This models the case when the knowledge of the frequency is not accurate. We also see in Figure 6 that the performances of both methods are the same with known or unknown  $\sigma^2$ , and probabilities of correct classification goes to 1 as  $\sigma^2 \rightarrow 0$ .



**Figure 6. Probability of Correct Classification for Both Methods.**

### 3.4. Robust Signal Detection Using the EEF

#### 3.4.1. Introduction

It is well known that in detection problems, the Neyman-Pearson (NP) decision rule is optimal in that it maximizes the probability of detection  $P_D$  for a given probability of false alarm  $P_{FA} = \alpha$  [37]. It decides  $H_1$  if

$$L(\mathbf{x}) = \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})} > \tau \quad (136)$$

where  $p_1(\mathbf{x})$  and  $p_0(\mathbf{x})$  are the probability density functions (PDFs) of the observed samples  $\mathbf{x}$  under  $H_1$  and  $H_0$  respectively, and  $\tau$  is the threshold given by

$$P_{FA} = \int_{\{\mathbf{x}: L(\mathbf{x}) > \tau\}} p_0(\mathbf{x}) d\mathbf{x} = \alpha \quad (137)$$

However, the NP rule requires complete knowledge of the PDF under each hypothesis. For many practical detection problems, the PDF under each hypothesis is parameterized by some unknown parameters. Due to the unknown parameters, the PDFs are not completely specified. Therefore, suboptimal detectors are implemented and the NP detector can be used as a performance upper bound. One popular suboptimal detector is the generalized likelihood ratio test (GLRT) [37], which replaces the unknown parameters by their maximum likelihood estimates (MLEs) [38]. It has been shown that the GLRT is asymptotically uniformly most powerful (UMP) among all tests that are invariant [39]. In some other problems, it is even difficult to obtain the parameterized PDFs. This may happen in the cases when we do not have enough training data to accurately estimate the PDF, or to find a parameterized model that fits the data, especially when dimension of the samples is high. Hence, sensor integration techniques have been used to construct a PDF based on all the available information. The copula based framework is used in [40], [41] to estimate the joint PDF from the marginal PDFs. Joint PDF construction from marginal PDFs is also considered in [42] based on the exponentially embedded family (EEF). The EEF is extended to the case when even the marginal PDFs are unknown [43].

We consider a robust detection problem where the alternative hypothesis  $H_1$  has multiple candidate models. Since the PDF under  $H_1$  is not completely known or even parameterized, the optimal NP detector or the suboptimal GLRT cannot be implemented. Because of the lack of information about the PDF under  $H_1$ , we apply the multimodal sensor integration method proposed in [43] to combine the multiple candidate models into a parameterized EEF, and use the EEF as the parameterized PDF under  $H_1$ . We further prove that the EEF is asymptotically optimal, as it will converge to the true candidate model as the number of samples goes to infinity.

## Problem Statement

Consider a hypothesis testing problem where the alternative hypothesis has two candidate models, i.e.,

$$\begin{aligned} H_0 : \mathbf{x} &\sim p_0(\mathbf{x}) \\ H_1 : \mathbf{x} &\sim p_1(\mathbf{x}) \text{ or } \mathbf{x} \sim p_2(\mathbf{x}) \end{aligned} \quad (138)$$

where  $\mathbf{x}$  is the  $N \times 1$  observed sample vector. So under  $H_1$ ,  $\mathbf{x}$  can be generated from either  $p_1$  or  $p_2$ . This is different from the traditional multiple hypothesis testing problem, where each hypothesis has only one candidate model. Note that we use two candidate models for simplicity, and the results can be easily extended to multiple candidate models. It is clear that if the true

PDF under  $H_1$  is  $p_1(\mathbf{x})$ , then the optimal NP detector decides  $H_1$  if  $T_1(\mathbf{x}) = \ln \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})} > \tau_1$  for some threshold  $\tau_1$ . Similarly, if the true PDF under  $H_1$  is  $p_2(\mathbf{x})$ , then the optimal NP detector decides  $H_1$  if  $T_2(\mathbf{x}) = \ln \frac{p_2(\mathbf{x})}{p_0(\mathbf{x})} > \tau_2$  for some threshold  $\tau_2$ . Unfortunately, we only know that the true PDF is either one of the two. If the wrong PDF has been used, it is expected that the detection performance will degrade. Therefore, we need to find a way to combine the candidate models using all the available information. In the next section, we will use the multimodal sensor integration technique to construct the parameterized PDF under  $H_1$ .

### 3.4.2. Sensor Integration and the EEF

It has been shown in [43] that if we observe measurements from two sensors  $\mathbf{T}_1(\mathbf{x})$  ( $p_1 \times 1$ ) and  $\mathbf{T}_2(\mathbf{x})$  ( $p_2 \times 1$ ) which are transformations of the sample vector  $\mathbf{x}$ , we can construct the PDF of  $\mathbf{x}$  under  $H_1$  as

$$p_{\boldsymbol{\eta}_1, \boldsymbol{\eta}_2}(\mathbf{x}) = \exp \left[ \boldsymbol{\eta}_1^T \mathbf{T}_1(\mathbf{x}) + \boldsymbol{\eta}_2^T \mathbf{T}_2(\mathbf{x}) - K(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2) + \ln p_0(\mathbf{x}) \right] \quad (139)$$

where  $K(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2) = \ln E_0 \left( \exp \left[ \boldsymbol{\eta}_1^T \mathbf{T}_1(\mathbf{x}) + \boldsymbol{\eta}_2^T \mathbf{T}_2(\mathbf{x}) \right] \right)$  is the cumulant generating function that normalizes the PDF to integrate to one. The subscript  $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2$  on  $p$  implies that the PDF is parameterized by  $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2$ .

In the problem discussed above, we are not directly given sensor outputs  $\mathbf{T}_1(\mathbf{x})$  and  $\mathbf{T}_2(\mathbf{x})$ . But we can consider  $T_1(\mathbf{x}) = \ln \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})}$  as the optimal detection statistic if  $p_1(\mathbf{x})$  is true, and  $T_2(\mathbf{x}) = \ln \frac{p_2(\mathbf{x})}{p_0(\mathbf{x})}$  as the optimal detection statistic if  $p_2(\mathbf{x})$  is true. Note that  $T_1(\mathbf{x})$  and  $T_2(\mathbf{x})$  are scalars here, so we do not use boldface. Hence,  $T_1(\mathbf{x}) = \ln \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})}$  and  $T_2(\mathbf{x}) = \ln \frac{p_2(\mathbf{x})}{p_0(\mathbf{x})}$  can be considered as the measurements from two sensors. Each sensor is optimal for its corresponding candidate hypothesis. Then we have the parameterized PDF of  $\mathbf{x}$  under  $H_1$  according to (139) as

$$\begin{aligned} p_{\eta_1, \eta_2}(\mathbf{x}) &= \exp \left[ \eta_1 T_1(\mathbf{x}) + \eta_2 T_2(\mathbf{x}) - K(\eta_1, \eta_2) + \ln p_0(\mathbf{x}) \right] \\ &= \exp \left[ \eta_1 \ln \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})} + \eta_2 \ln \frac{p_2(\mathbf{x})}{p_0(\mathbf{x})} - K(\eta_1, \eta_2) + \ln p_0(\mathbf{x}) \right] \end{aligned} \quad (140)$$

where  $K(\eta_1, \eta_2) = \ln E_0 \left( \exp \left[ \eta_1^T T_1(\mathbf{x}) + \eta_2^T T_2(\mathbf{x}) \right] \right)$ . Note that  $K(1, 0) = \ln \int p_1(\mathbf{x}) d\mathbf{x} = 0$ . As a result, when  $\eta_1 = 1$  and  $\eta_2 = 0$ , we have  $p_{1,0}(\mathbf{x}) = p_1(\mathbf{x})$ . Similarly,  $p_{0,1}(\mathbf{x}) = p_2(\mathbf{x})$ .

Since  $p_{\eta_1, \eta_2}(\mathbf{x})$  is the parameterized PDF under  $H_1$ , now we can use the GLRT for detection. The GLRT requires the MLEs for  $\eta_1$  and  $\eta_2$ . The MLEs are obtained by maximizing the log-likelihood function  $\ln p_{\eta_1, \eta_2} = \eta_1 T_1(\mathbf{x}) + \eta_2 T_2(\mathbf{x}) - K(\eta_1, \eta_2) + \ln p_0(\mathbf{x})$  or equivalently maximizing  $\eta_1 T_1(\mathbf{x}) + \eta_2 T_2(\mathbf{x}) - K(\eta_1, \eta_2)$ . Since  $K(\eta_1, \eta_2)$  is a convex function by Holder's inequality [44], the MLEs can be solved by taking the derivatives with respect to  $\eta_1$  and  $\eta_2$ , and setting them to zeros. That is, the MLEs for  $\eta_1$  and  $\eta_2$  should satisfy

$$\ln \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})} = \frac{\partial K(\eta_1, \eta_2)}{\partial \eta_1} \quad (141)$$

$$\ln \frac{p_2(\mathbf{x})}{p_0(\mathbf{x})} = \frac{\partial K(\eta_1, \eta_2)}{\partial \eta_2} \quad (142)$$

Once the MLEs  $\hat{\eta}_1$  and  $\hat{\eta}_2$  are solved from (141), the GLRT decides  $H_1$  if

$$\ln \frac{p_{\hat{\eta}_1, \hat{\eta}_2}(\mathbf{x})}{p_0(\mathbf{x})} = \hat{\eta}_1 T_1(\mathbf{x}) + \hat{\eta}_2 T_2(\mathbf{x}) - K(\hat{\eta}_1, \hat{\eta}_2) > \tau \quad (143)$$

### 3.4.3. Asymptotic Optimality of the EEF

In sensor integration, the EEF has been shown to be asymptotically optimal in that it minimizes the Kullback-Leibler (KL) divergence between the true PDF and the parameterized PDF [43]. The KL divergence is known as a non-symmetric measure of difference between two PDFs. For two PDFs  $p_1$  and  $p_0$ , it is defined as

$$D(p_1 \| p_0) = \int p_1(\mathbf{x}) \ln \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})} d\mathbf{x} \quad (144)$$

The KL divergence is non-negative or  $D(p_1 \| p_0) \geq 0$  with equality if and only if  $p_1 = p_0$  [9]. It also measures the asymptotic performance for detection by Stein's lemma [46].

In this section, we will also show that the EEF is asymptotically optimal in that  $p_{\hat{\eta}_1, \hat{\eta}_2}(\mathbf{x})$  converges to the true PDF. Note that we have shown that when  $\eta_1 = 1$  and  $\eta_2 = 0$ ,  $p_{\eta_1, \eta_2}(\mathbf{x}) = p_1(\mathbf{x})$ . Hence, it is expected that if the true PDF is  $p_1(\mathbf{x})$ ,  $\hat{\eta}_1$  should be close to 1 and  $\hat{\eta}_2$  be close to 0. This is true, as we will prove next, that when the true PDF is  $p_1$ ,  $\hat{\eta}_1 \xrightarrow{P} 1$  and  $\hat{\eta}_2 \xrightarrow{P} 0$  as the number of samples  $L \rightarrow \infty$  where  $\xrightarrow{P}$  denotes convergence in probability.



Assume that we have  $L$  independent identically distributed (IID) sample vectors  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{L-1}$ . The parameterized PDF in (140) can be easily extended as

$$\begin{aligned} p_{\eta_1, \eta_2}(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{L-1}) &= \exp \left[ \eta_1 \sum_{i=0}^{L-1} T_1(\mathbf{x}_i) + \eta_2 \sum_{i=0}^{L-1} T_2(\mathbf{x}_i) - LK(\eta_1, \eta_2) + \sum_{i=0}^{L-1} \ln p(\mathbf{x}_i; H_0) \right] \\ &= \exp \left[ \eta_1 \sum_{i=0}^{L-1} \ln \frac{p_1(\mathbf{x}_i)}{p_0(\mathbf{x}_i)} + \eta_2 \sum_{i=0}^{L-1} \ln \frac{p_2(\mathbf{x}_i)}{p_0(\mathbf{x}_i)} - LK(\eta_1, \eta_2) + \sum_{i=0}^{L-1} \ln p(\mathbf{x}_i; H_0) \right] \end{aligned} \quad (145)$$

The MLEs  $\eta_1$  and  $\eta_2$  are found by solving

$$\frac{1}{L} \sum_{i=0}^{L-1} \ln \frac{p_1(\mathbf{x}_i)}{p_0(\mathbf{x}_i)} = \frac{\partial K(\eta_1, \eta_2)}{\partial \eta_1} \quad (146)$$

$$\frac{1}{L} \sum_{i=0}^{L-1} \ln \frac{p_2(\mathbf{x}_i)}{p_0(\mathbf{x}_i)} = \frac{\partial K(\eta_1, \eta_2)}{\partial \eta_2} \quad (147)$$

Assume that  $T_1(\mathbf{x}_i)$  and  $T_2(\mathbf{x}_i)$  are linearly independent [47], then  $\eta_1$  and  $\eta_2$  are identifiable, and hence  $K(\eta_1, \eta_2)$  is strictly convex [48]. Then the solution of (145, 146) is unique. Let

$$\mathbf{t} = \begin{bmatrix} \frac{1}{L} \sum_{i=0}^{L-1} \ln \frac{p_1(\mathbf{x}_i)}{p_0(\mathbf{x}_i)} \\ \frac{1}{L} \sum_{i=0}^{L-1} \ln \frac{p_2(\mathbf{x}_i)}{p_0(\mathbf{x}_i)} \end{bmatrix} \quad (148)$$

$$\boldsymbol{\eta} = \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \quad (149)$$

and

$$\mathbf{f}(\boldsymbol{\eta}) = \begin{bmatrix} \frac{\partial K(\eta_1, \eta_2)}{\partial \eta_1} \\ \frac{\partial K(\eta_1, \eta_2)}{\partial \eta_2} \end{bmatrix} \quad (150)$$

Then (145, 146) can be written as

$$\mathbf{f}(\boldsymbol{\eta}) = \mathbf{t} \quad (151)$$

Since for each  $\mathbf{t}$ , the solution of  $\boldsymbol{\eta}$  is unique, there is a one-to-one correspondence between  $\mathbf{t}$  and  $\boldsymbol{\eta}$ . Hence, we have the inverse function

$$\boldsymbol{\eta} = \mathbf{f}^{-1}(\mathbf{t}) \quad (152)$$

Assume that the true PDF is  $p_1(\mathbf{x})$ , then we have

$$\mathbf{t} = \begin{bmatrix} \frac{1}{L} \sum_{i=0}^{L-1} \ln \frac{p_1(\mathbf{x}_i)}{p_0(\mathbf{x}_i)} \\ \frac{1}{L} \sum_{i=0}^{L-1} \ln \frac{p_2(\mathbf{x}_i)}{p_0(\mathbf{x}_i)} \end{bmatrix} \xrightarrow{P} \begin{bmatrix} E_1 \left( \ln \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})} \right) \\ E_1 \left( \ln \frac{p_2(\mathbf{x})}{p_0(\mathbf{x})} \right) \end{bmatrix} = \mathbf{t}^* \quad (153)$$

as  $L \rightarrow \infty$  by the law of large numbers. Let  $\boldsymbol{\eta}^* = [10]^T$ . Since

$$\begin{bmatrix} \frac{\partial K(\eta_1, \eta_2)}{\partial \eta_1} \\ \frac{\partial K(\eta_1, \eta_2)}{\partial \eta_2} \end{bmatrix} = \begin{bmatrix} E \left( \ln \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})} \right) \\ E \left( \ln \frac{p_2(\mathbf{x})}{p_0(\mathbf{x})} \right) \end{bmatrix} \quad (154)$$

we have

$$\mathbf{f}(\boldsymbol{\eta}^*) = \begin{bmatrix} \frac{\partial K(\eta_1, \eta_2)}{\partial \eta_1} \\ \frac{\partial K(\eta_1, \eta_2)}{\partial \eta_2} \end{bmatrix}_{\boldsymbol{\eta}=\boldsymbol{\eta}^*} = \begin{bmatrix} E_* \left( \ln \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})} \right) \\ E_* \left( \ln \frac{p_2(\mathbf{x})}{p_0(\mathbf{x})} \right) \end{bmatrix} = \begin{bmatrix} E_1 \left( \ln \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})} \right) \\ E_1 \left( \ln \frac{p_2(\mathbf{x})}{p_0(\mathbf{x})} \right) \end{bmatrix} = \mathbf{t}^* \quad (155)$$

Note that the expected value with respect to  $p_*$  is the same as that with respect to  $p_1$ . This is because as we have shown that, when  $\boldsymbol{\eta}^* = [\eta_1^* \eta_2^*]^T = [10]^T$ ,  $p_{\boldsymbol{\eta}^*}(\mathbf{x}) = p_1(\mathbf{x})$ . From (154), we have

$$\boldsymbol{\eta}^* = \mathbf{f}^{-1}(\mathbf{t}^*) \quad (156)$$

We assume that  $\mathbf{f}^{-1}$  is continuous. From (152)  $\mathbf{t} \xrightarrow{P} \mathbf{t}^*$ , then we have [49]

$$\boldsymbol{\eta} = \mathbf{f}^{-1}(\mathbf{t}) \xrightarrow{P} \mathbf{f}^{-1}(\mathbf{t}^*) = \boldsymbol{\eta}^* = [10]^T \quad (157)$$

It shows that  $\hat{\eta}_1 \xrightarrow{P} 1$  and  $\hat{\eta}_2 \xrightarrow{P} 0$  as the number of samples  $L \rightarrow \infty$  if the true PDF is  $p_1$ . This also implies the asymptotic optimality of the EEF since the parameterized PDF  $p_{\hat{\eta}_1, \hat{\eta}_2}$  converges to the true PDF  $p_1$  as  $L \rightarrow \infty$ . From this analysis, we can use  $\hat{\eta}_1$  and  $\hat{\eta}_2$  as indicator of the true PDF.

#### 3.4.4. Example: Detection Problem with Sensor Integration

Consider a detection problem with

$$\begin{aligned} H_0 : \mathbf{x} &= \mathbf{w} \\ H_1 : \mathbf{x} &= \mathbf{A} + \mathbf{w} \text{ or } \mathbf{x} = \mathbf{s} + \mathbf{w} \end{aligned} \quad (158)$$

where  $\mathbf{x}$  is the  $N \times 1$  sample vector,  $\mathbf{A} = [A \ A \ \dots \ A]^T$  is completely known,  $\mathbf{s} \sim N(0, \sigma_s^2 \mathbf{I})$ ,  $\mathbf{w} \sim N(0, \sigma^2 \mathbf{I})$  with  $\sigma_s^2, \sigma^2$  known. If the candidate model  $\mathbf{x} = \mathbf{A} + \mathbf{w}$  is used under  $H_1$ , the optimal NP detector decides  $H_1$  if

$$T_1(\mathbf{x}) = \ln \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})} = \frac{2\mathbf{A}^T \mathbf{x} - \mathbf{A}^T \mathbf{A}}{2\sigma^2} = \frac{\mathbf{A}^T \mathbf{x}}{\sigma^2} + c_1 > \tau_1 \quad (159)$$

with  $c_1$  a constant that does not depend on  $\mathbf{x}$ . This is also known as the matched filter. If the candidate model  $\mathbf{x} = \mathbf{s} + \mathbf{w}$  is used under  $H_1$ , the optimal NP detector decides  $H_1$  if

$$T_2(\mathbf{x}) = \ln \frac{p_2(\mathbf{x})}{p_0(\mathbf{x})} = \left( \frac{1}{2\sigma^2} - \frac{1}{2(\sigma_s^2 + \sigma^2)} \right) \mathbf{x}^T \mathbf{x} + c_2 = \frac{\sigma_s^2 \mathbf{x}^T \mathbf{x}}{2\sigma^2(\sigma_s^2 + \sigma^2)} + c_2 > \tau_2 \quad (160)$$

with  $c_2$  a constant. This is equivalent to an energy detector.

The parameterized PDF constructed according to (140) is

$$p_{\eta_1, \eta_2}(\mathbf{x}) = \exp[\eta_1 T_1(\mathbf{x}) + \eta_2 T_2(\mathbf{x}) - K(\eta_1, \eta_2) + \ln p(\mathbf{x}; H_0)] \quad (161)$$

where  $\ln p_0(\mathbf{x}) = -\frac{\mathbf{x}^T \mathbf{x}}{2\sigma^2} + c_0$  with  $c_0$  a constant. The constants in  $T_1(\mathbf{x})$ ,  $T_2(\mathbf{x})$  and  $\ln p_0(\mathbf{x})$  can be ignored as  $K(\eta_1, \eta_2)$  will normalize the PDF. Hence, the EEF is

$$\begin{aligned}
& p_{\eta_1, \eta_2}(\mathbf{x}) \\
&= \exp \left[ \eta_1 \frac{\mathbf{A}^T \mathbf{x}}{\sigma^2} + \eta_2 \frac{\sigma_s^2 \mathbf{x}^T \mathbf{x}}{2\sigma^2(\sigma_s^2 + \sigma^2)} - K(\eta_1, \eta_2) - \frac{\mathbf{x}^T \mathbf{x}}{2\sigma^2} \right] \\
&= \exp \left[ - \underbrace{\left( \frac{1}{2\sigma^2} - \frac{\eta_2 \sigma_s^2}{2\sigma^2(\sigma_s^2 + \sigma^2)} \right)}_a \mathbf{x}^T \mathbf{x} + \eta_1 \frac{\mathbf{A}^T \mathbf{x}}{\sigma^2} - K(\eta_1, \eta_2) \right] \\
&= \exp \left[ -a \left\| \mathbf{x} - \frac{\eta_1 \mathbf{A}}{2a\sigma^2} \right\|^2 + \frac{\eta_1^2 \mathbf{A}^T \mathbf{A}}{4a\sigma^4} - K(\eta_1, \eta_2) \right] \\
&= \frac{1}{(\pi/a)^{N/2}} \exp \left[ -a \left\| \mathbf{x} - \frac{\eta_1 \mathbf{A}}{2a\sigma^2} \right\|^2 \right]
\end{aligned} \tag{162}$$

We assume that  $0 \leq \eta_1, \eta_2 \leq 1$ . The MLE is found by maximizing

$$\ln p_{\eta_1, \eta_2}(\mathbf{x}) = -\frac{N}{2} \ln(\pi/a) - a \left\| \mathbf{x} - \frac{\eta_1 \mathbf{A}}{2a\sigma^2} \right\|^2 \tag{163}$$

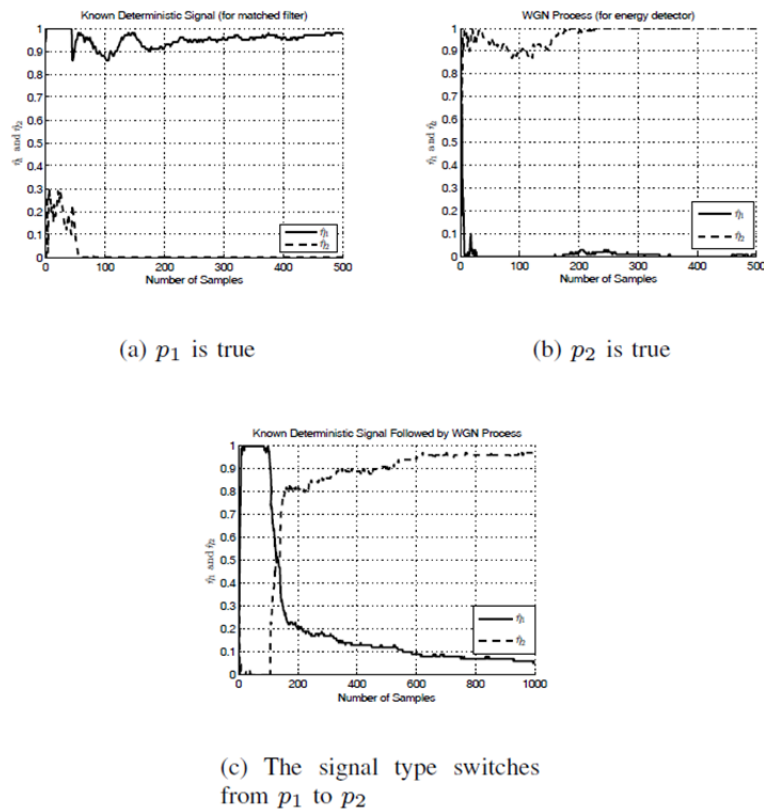
over  $0 \leq \eta_1, \eta_2 \leq 1$  where  $a = \left( \frac{1}{2\sigma^2} - \frac{\eta_2 \sigma_s^2}{2\sigma^2(\sigma_s^2 + \sigma^2)} \right) = \frac{1}{2\sigma^2} \frac{(1-\eta_2)\sigma_s^2 + \sigma^2}{\sigma_s^2 + \sigma^2}$ . Then we can decide  $H_1$  if

$$\ln \frac{p_{\hat{\eta}_1, \hat{\eta}_2}(\mathbf{x})}{p_0(\mathbf{x})} > \tau \tag{164}$$

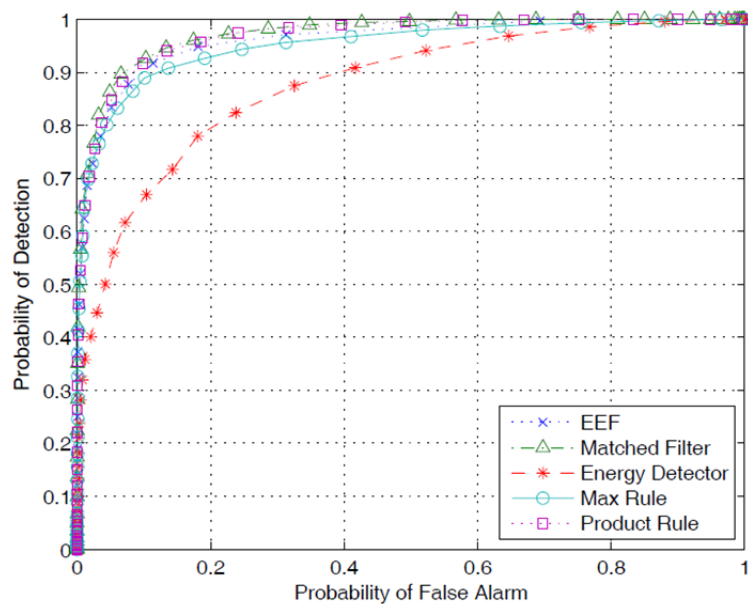
### 3.4.5. Simulation Results

For the example above, we let  $A = 1.2$ ,  $\sigma_s^2 = 2$ , and  $\sigma^2 = 1$ . In Figure 7(a), we plot  $\hat{\eta}_1$  and  $\hat{\eta}_2$  as  $N$  increases when all the samples are generated from  $p_1$  or  $\mathbf{x} = \mathbf{A} + \mathbf{w}$ . We see that  $\hat{\eta}_1 \rightarrow 1$  and  $\hat{\eta}_2 \rightarrow 0$  as  $N \rightarrow \infty$  which corresponds to  $p_1$ . For Figure 7(b), all the samples are generated from  $p_2$  or  $\mathbf{x} = \mathbf{s} + \mathbf{w}$  and we can see that  $\hat{\eta}_1 \rightarrow 0$  and  $\hat{\eta}_2 \rightarrow 0$  as  $N \rightarrow \infty$  which corresponds to  $p_2$ . In Figure 7(c), the first 100 samples are generated from  $p_1$  and the next 1900 samples are generated from  $p_2$ . Therefore,  $\hat{\eta}_1$  and  $\hat{\eta}_2$  can be used to track the models. Then we plot the receiver operating characteristic (ROC) curves for the EEF, matched filter, energy detector, max rule, product rule [50] when the true PDF is  $p_1$  in Figure 8 (using  $N = 5$ ) and the ROC curves for these methods when  $p_2$  is true in Figure 9 (using  $N = 25$ ). Note that the matched filter is the optimal NP detector in Figure 8 and the energy detector is the optimal NP detector in Figure 9.

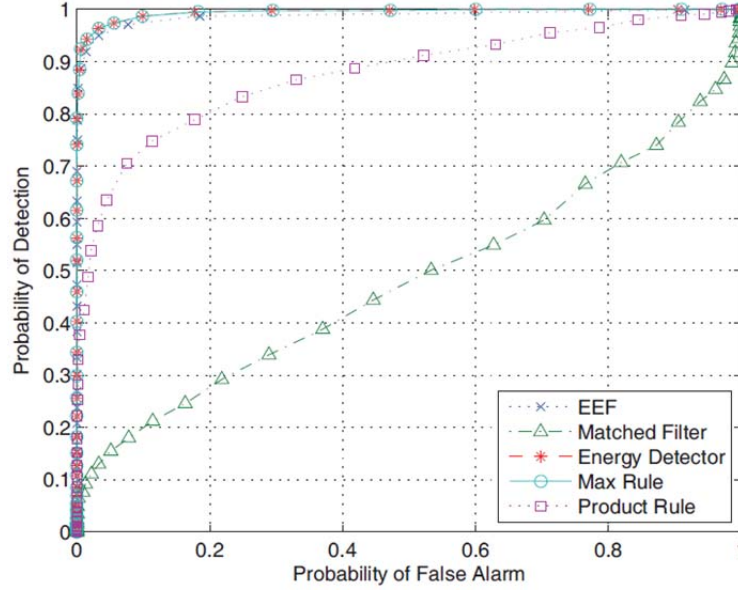
We can see that the performance of the EEF is close to the NP detector in both cases, and therefore is more robust.



**Figure 7. PDF Parameters for Different Models.**



**Figure 8. ROC curves for different detectors when  $\mathbf{x} = \mathbf{A} + \mathbf{w}$ .**



**Figure 9. ROC Curves for Different Detectors when  $\mathbf{x} = \mathbf{s} + \mathbf{w}$ .**

### 3.5. Conclusions

A novel method of constructing the joint PDF of the measurements from a distributed multiple sensor systems has been proposed. Only a reference PDF is needed in the construction. It is shown that moment matching and MLE are equivalent. The performance of our method has shown to be as good as the clairvoyant GLRT and estimated MAP classifier respectively for detection and classification, while less information is needed for our method.

The EEF has been used for a robust detection problem where the alternative hypothesis  $H_1$  has multiple candidate models. A parameterized PDF under  $H_1$  is constructed based on the EEF. The asymptotic optimality of the EEF is proved. The EEF and existing classifier combining rules are compared in simulations, and it shows that the EEF has comparable performance with the optimal NP detector.

We have shown the feasibility and benefits of using the EEF for detection and classification problems where the PDF of the sensor measurements is unknown. The EEF has important applications in multi-modal radar systems since it allows us to combine data of different modality (for example, MTI data and EO data) from a statistical standpoint. It is constructed based on the measurements from different sensors, and therefore, it is robust in practice where we do not have the exact signal model. A global decision is made by the central processor which combines all the available information from different sensors. The EEF is computationally efficient because it belongs to the exponential family and the maximum likelihood estimation is a convex optimization problem. The EEF asymptotically minimizes the Kullback-Leibler divergence between the true PDF and the exponential family. As a result, the performance of the detector or classifier is expected to be asymptotically optimal.

## 4. Algorithm Development and Implementation

We address an important sub-problem of using multiple, spatially distributed, adaptive multi-modal sensors for multiple target tracking and data association. When left unmodeled or ignored, inherent field-of-view (FOV) limitations for each mode and sensor present a challenge to accurate tracking, appropriate sensor management, and system robustness. This is especially true in the context of increasing sensor adaptivity, which may naively sacrifice multiple target performance for single-target hyper accuracy. In light of these limitations, we propose a relatively simple variance-penalty oriented modeling solution that effectively presents the FOV as a new design parameter. A novel closed-loop adaptive mode-parameter selection algorithm incorporating this penalty model is also proposed for use in both ideal and cluttered sensing environments. Simulations demonstrate the necessity and benefits of the proposed model and algorithm in a variety of tracking and association contexts as compared to completely unmodeled/non-adaptive methods.

### 4.1. Methods, Assumptions, and Procedures

The purpose of this section is to summarize the research effort into limited field-of-view (FOV) effects in multi-modal sensing applications.

The role of multimodal sensing in multiple target tracking and data fusion has received increased attention and research efforts, evidenced by the increasing available literature on the topic and the existence of coordinated research efforts like the one this report is based on. By combining multiple sensor types into a single integrated system, the object is to obtain superior performance compared with employing the modes individually. A commonly investigated sensor grouping is the combining of active and passive sensor modes, exemplified by the radio-frequency radar/electro-optical imaging (RF/EO) sensor pairing. It has been shown by Blasch and Yang in [51] that this pairing results in improved data fusion, reducing misassociation and tracking errors in environments with clutter. These systems have also been shown to benefit from adaptive mode and parameter selection, which forms the basis for the rest of the following discussion.

Adaptive methods for active sensing modes have been extensively studied in a variety of contexts. A notable example are the pioneering works by Kershaw and Evans in [52], [53] on radar waveform adaptation for single target tracking in ideal and cluttered sensing. However, both of these works were limited to single-target tracking and restricted sets of transmit waveform patterns, which greatly simplified the required optimization scheme. More modern examples of radar adaptivity for tracking and association improvement can be found in the review paper by Sira, *et al.* [54] and the references within. Adaptation for passive sensor systems has also been considered, *e.g.* [62].

More recently, work has focused on systems fusing adaptivity and multimodal sensing for enhanced multitarget tracking performance. The work by Zhang, *et al.* [55] exemplifies this effort, considering a system with multiple joint radar/electro-optical sensors that are independently tuned throughout the tracking period. This is accomplished through a waveform-agile particle filter to track multiple targets in clutter, using heuristic methods to select the system parameter set that minimized the posterior Cramer-Rao lower bound (PCRLB).

In order to appropriately characterize the benefits of adaptivity in any system, much less ones exploiting multiple sensing modes simultaneously, attention must be paid to constraints that limit the extent of the adaptation. By appropriately modeling realistic constraints, the impact of a given method in an actual deployment scenario can be better assessed, and modifications can be tailored to mitigate or even exploit the constraint's effect. The literature provides a variety of analyses & algorithms focused on constraint modeling in sensor management & target tracking, including, for example, single-sensor target occlusion [56] and energy conservation in the context of sleeping policies [57].

However, of the aforementioned existing literature, however, almost none in the signal processing literature consider the effects of finite sensor field-of-view on tracking performance or system adaptability. This effect is particularly pronounced in EO sensors, since a target outside the FOV is essentially unmeasurable. For the RF modes, out-of-“view” measurements take the form of range or doppler aliased returns, if they are even validated at all. Due to the more noticeable characterization of this effect in optical systems, the concept of field-of-view limitations has been mostly limited to the computer vision and camera literature, *e.g.* [58–60]. Furthermore, most work in the signal processing community has concentrated on controlling precise aspects of (usually active) sensors, like low-level waveform or beamform design. Our consideration of multimodal sensing, however, necessitates a system-level perspective on “waveform”-adaptivity in the spirit of Daum [61] nearly twenty years ago, which emphasized the impact of resolution on adaptive parameter schemes for multitarget tracking. In this report, we take a similar approach to the FOV problem.

Our contribution during this performance period has been to develop a novel approach to this fundamental limitation. Using penalty functions that depend on the mode, measurement function, and control variables, we have parameterized the FOV effect in a mathematically simple, but powerful manner. These penalty functions augment known models of each measurement's noise variance (itself a function of the control) by assigning significantly larger variances to targets outside or nearly outside the FOV. This formulation permits a combined covariance compensation & joint optimization scheme to natively account for the strengths and limitations of each mode.

## 4.2. Target Dynamic Models

In this work, we consider a common non-maneuvering point-target dynamic model for the targets we wish to track, operating in discrete time with a sampling period of  $T$  seconds. For simplicity, we only consider two-dimensional movement in the  $(x, y)$  plane. At time step  $k$ , the state vector for the  $i$ th target is comprised of Cartesian position and velocity coordinates:

$$\mathbf{x}_k^i = [x_k^i \quad \dot{x}_k^i \quad y_k^i \quad \dot{y}_k^i]^T \quad (165)$$

where the superscript  $i$  denotes a target index,

Assuming a linear (or sufficiently linearized) target motion model, the state vector for a single target evolves in time as

$$\mathbf{x}_{k+1}^i = \mathbf{A}_k^i \mathbf{x}_k^i + \mathbf{w}_k^i \quad (166)$$



where  $\mathbf{A}_k^i$  is the maneuver-dependent system matrix and  $\mathbf{w}_k^i$  is zero-mean Gaussian process noise with covariance matrix  $\mathbf{Q}_k^i$ .

If we assume that there are  $N_{T,k}$  targets present in the greater surveillance area at time  $k$ , then the multitarget state is the set  $\mathbf{x}_k = \{\mathbf{x}_k^i\}_{i=1}^{N_{T,k}}$ . For this section, we assume that  $N_{T,k}$  is known *a priori* and remains constant for all values of  $k$ . With this simplification, the set becomes a single multitarget supervector formed from stacking the per-target state vectors in index order, or  $\mathbf{x}_k = [\mathbf{x}_k^{(1)T} \dots \mathbf{x}_k^{(N_{T,k})T}]^T$ . This multitarget state vector evolves as

$$\mathbf{x}_{k+1} = \mathcal{A}_k \mathbf{x}_k + \mathbf{w}_{k+1} \quad (167)$$

where  $\mathcal{A}_k = \text{blockdiag}[\mathbf{A}_k^{(1)}, \dots, \mathbf{A}_k^{(N_{T,k})}]$  and  $\mathbf{w}_{k+1}$  is a supervector of per-target process noise components constructed similarly to the state supervector with a covariance of  $\mathcal{Q}_k = \text{blockdiag}[\mathbf{Q}_k^{(1)}, \dots, \mathbf{Q}_k^{(N_{T,k})}]$ , provided there is no coupling of the target states.

### 4.3. Sensor Measurement Models

This subsection describes the measurement models for the multi-modal sensors considered in our analysis. We first describe some general characteristics common to the rest of the section. Suppose we observe the aforementioned target environment with  $N_S$  homogenous multimodal sensors distributed in the surveillance region, with each sensor  $p$  comprising of  $M$  modes. Furthermore, we can parameterize the  $(p, s)$ th sensor/mode's operation at time  $k$  by the vector  $\mathbf{p}_{(p,s),k}$ . In the interests of notational simplicity, the paired sensor-mode indexing  $(p, s)$  is collapsed into a single (pseudo-)sensor index  $j$  defined by the mapping  $j = M(p - 1) + s$ . The work discussed in this report examined finite FOV effects in both ideal (perfect detection) and cluttered sensing environments; as such, the general assumptions and models for both cases will be discussed in the subsections below.

#### 4.3.1. Ideal Sensing Environment

We begin our sensor modeling analysis by considering the idealized scenario, where sensor/modes are assumed to exist in a non-cluttered environment, have perfect detection characteristics, and always correctly associate measurements with their targets. This scenario permits us to construct our proposed method free of other concerns.

Under these assumptions, the  $j$ th sensor/mode measures the  $i$ th target as

$$\mathbf{y}_{j,k}^i = \mathbf{c}_j(\mathbf{x}_k^i) + \mathbf{v}_{j,k}^i \quad (168)$$

where  $\mathbf{c}_j(\mathbf{x}_k^i)$  is the (possibly non-linear) measurement function and  $\mathbf{v}_{j,k}^i$  is measurement noise. These measurement noise sequences are assumed to be zero-mean Gaussian with system parameter-dependent covariance  $\mathbf{R}_j^i(\mathbf{p}_{j,k})$  that are uncorrelated with both the target process noise and the measurement noise from any other sensor/mode.

For the desired & examined RF/EO mode pairing, the measurement function in two dimensions corresponds to the RF sensor measuring range and range-rate and the EO sensor measuring azimuth. The RF sensors obtain their respective measurements by matched filtering of target returns in a range-doppler, resulting in range and range-rate estimates for each target. For the  $j$ th sensor located at coordinates  $(\bar{x}_j, \bar{y}_j)$ , the range  $r_j^i$  to and the range-rate  $\dot{r}_j^i$  of the  $i$ th target are

$$r_j^i = \sqrt{(x^i - \bar{x}_j)^2 + (y^i - \bar{y}_j)^2} \quad (169)$$

$$\dot{r}_j^i = \frac{\dot{x}^i(x^i - \bar{x}_j) - \dot{y}^i(y^i - \bar{y}_j)}{r_j^i}. \quad (170)$$

The EO sensors obtain (in this particular scenario) azimuth measurements from their captured image. Hence, the EO measurement equation is  $\eta = \arcsin\left(\frac{(x^i - x_{p,j})}{r_j^i}\right)$ . For one sensor, then, the overall measurement vector for the  $i$ th target at time step  $k$  is  $\mathbf{c}_{j,k}^i = [r_j^i \quad \dot{r}_j^i \quad \eta_j^i]^T$ . Each of these measurements is corrupted by zero-mean Gaussian noise with a variance that depends on each sensor's measurement parameters, which can be collected in the parameter vector  $\mathbf{p}_{j,k}$ . If we assume that the measurements are uncorrelated from sensor to sensor, and further, that multiple measurements from a single sensor are also uncorrelated, then the measurement covariance matrix for each target/sensor pair can be written as  $\mathbf{R}_{j,k}^i = \text{diag}(\sigma_r^2(\mathbf{p}_{j,k}), \sigma_{\dot{r}}^2(\mathbf{p}_{j,k}), \sigma_\eta^2(\mathbf{p}_{j,k}))$ , where the  $\text{diag}(\cdot)$  operator forms a matrix of its elements along the diagonal. Assuming that the RF mode is parameterized by a per-pulse bandwidth  $B_j$  and a pulse repetition frequency  $PRF_j$ , the measurement variances for range and range rate are given by  $\sigma_r^2(\mathbf{p}_{j,k}) = \frac{1}{12}(c/2B_j)^2$  and  $\sigma_{\dot{r}}^2(\mathbf{p}_{j,k}) = \frac{1}{12}(\lambda PRF_j/2N_{FFT})^2$  respectively, where  $c$  is the speed of light,  $\lambda$  is the carrier wavelength, and  $N_{FFT}$  is the Doppler integration length (FFT size) [62]. If we assume the mode captures a one-dimensional image  $N_{pix}$  pixels wide with a detector/pixel angular subtense of  $\alpha$ , the noise variance of the azimuth measurement is  $\sigma_\eta^2(\mathbf{p}_{j,k}) = \alpha^2/12$  [63]. For each of these measurement variances, we have assumed that the measurement noise is uniformly distributed in the given resolution cell/pixel and that only one target return originates from said resolution cell.

Temporarily dropping the time index for notational ease, we form a stacked vector of the measurements from all sensors for each target as  $\mathbf{y}^{(i)} = [\mathbf{y}_1^{(i)T} \cdots \mathbf{y}_{N_p}^{(i)T}]^T$ , and then concatenate these per-target vectors into a supervector  $\mathcal{Y} = [\mathbf{y}^{(1)T} \cdots \mathbf{y}^{(N_{T,k})T}]^T$ . This transforms the state-space measurement equation in Eq. 6 into:

$$\mathcal{Y}_k = \mathbf{c}(\chi_k) + \mathbf{v}_k \quad (171)$$

where  $\mathbf{c}(\chi_k)$  is the non-linear measurement function composed of the per-target, per-sensor measurement functions  $\mathbf{c}_j^i(\mathbf{x}_k^i)$  stacked as above. Similarly,  $\mathbf{v}_k$  is the identically generated noise

vector with block-diagonal covariance  $\mathcal{R}(\mathbf{p})$  (constructed from concatenating the measurement covariances), where  $\mathbf{p}$  is the overall system parameter vector composed of the per-sensor parameter vectors  $\mathbf{p}_j$ .

#### 4.3.2. Modifications for Cluttered Sensing Scenarios

Unfortunately, real life deployment situations use sensors with imperfect detection in environments with multiple non-target entities. Due to the inherent characteristics of each sensor/mode, the tracking system may reject true target returns and accept false returns as valid measurements. In addition, even if a target's return is correctly detected, there still exists ambiguity in regards to which target said measurement corresponds because of the potential for other targets and non-target returns. As a result, the idealized measurement model of the previous subsection is problematic and requires modification in order to account for these realities. This subsection describes these modifications.

Assume each sensor/mode  $j$  produces a number of raw returns that are preprocessed in a manner that produces kinematic measurements based on the target state vectors and clutter, with target detection probability  $P_{D,j}$  and false alarm rate  $P_{FA,j}$ . After initial detection, these returns are run through a validation procedure that compares them to a predicted measurement of each target (generated by the state estimation procedure described in Section 4.5 below). A measurement is considered valid if it falls within a predefined validation region centered on the predicted measurement. In this case, we consider a validation region given by the g-sigma ellipsoid

At time  $k$ ,  $m_{j,k}$  detected and preliminarily validated measurement vectors are obtained from sensor  $j$ , denoted by the set  $\mathcal{Y}_{j,k} = \{\mathbf{y}_{l,j,k}\}_{l=1}^{m_{j,k}}$ . If the  $l_j$ th measurement originated from the  $i$ th target, then  $\mathbf{y}_{l_j,j,k}$  is given by:

$$\mathbf{y}_{l_j,j,k}^i = \mathbf{c}_j(\mathbf{x}_k^i) + \mathbf{v}_{j,k}^i \quad (172)$$

where  $\mathbf{c}_j(\cdot)$  is the (possibly non-linear) mode measurement function and  $\mathbf{v}_{j,k}^i$  is a zero-mean Gaussian random vector denoting measurement noise with system parameter-dependent covariance  $\mathbf{R}_j(\mathbf{p}_{j,k})$ . We assume these noise sequences are uncorrelated with both the target process noise and the measurement noise from any other sensors or targets. The form of the mode covariance assumed under the clutter scenario is similar to the one considered above, but separated by mode (hence, the RF mode variances are separate from the EO mode variances, etc.).

Measurements that are not target-originated are assumed to originate from false alarms/clutter. These returns are assumed to be uniformly distributed throughout the  $j$ th sensor's validation region with volume  $V_{j,k}$  at time  $k$  -- that is, the probability distribution function for a clutter measurement  $\bar{\mathbf{v}}_{j,k}$  is  $p(\bar{\mathbf{v}}_{j,k}) = 1/V_{j,k}$ . The number of false alarms within a given validation region is a random variable modeled by a prior that depends on the available information about the clutter environment. In our case, the per-sensor clutter density  $\bar{\lambda}_j$  depends on the given false alarm rate  $P_{FA,j}$  and the size of the resolution cells (which vary depending on the sensor parameters). Since the clutter density is thus known, the number of false alarms in the volume is

modeled as a Poisson distributed random variable with parameter  $\bar{\lambda}_j V_{j,k}$ . If  $\bar{\lambda}_j$  was unknown, the false alarm count could be modeled as a diffuse prior, whose value is irrelevant since it cancels out of the probability distributions required in the state estimation.

#### 4.4. Ideal Sensing Environment: The Extended Kalman Filter

Under the perfect detection/association assumption, the natural state estimation choice, due to the nonlinear measurement equations, is the ubiquitous extended Kalman filter (EKF). While more sophisticated state estimation techniques have been used in the literature, the use of the EKF here is primarily to provide a simpler metric formulation for the optimization problem in Section 4.7.2. Using the multitarget state vector  $\mathbf{x}_k$  and given the measurement supervector  $\mathbf{y}_k$ , the EKF runs as follows:

*Prediction step:*

$$\hat{\mathbf{x}}_{k|k-1} = \mathcal{A}_k \hat{\mathbf{x}}_{k-1|k-1} \quad (173)$$

$$\mathcal{P}_{k|k-1} = \mathcal{A}_k \mathcal{P}_{k-1|k-1} \mathcal{A}_k^T + \mathcal{Q}_k \quad (174)$$

*Update step:*

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathcal{K}_k \mathbf{e}_k \quad (175)$$

$$\mathcal{P}_{k|k} = [\mathcal{P}_{k|k-1}^{-1} + \mathcal{C}_k^T \mathcal{R}^{-1}(\mathbf{p}_k) \mathcal{C}_k]^{-1} \quad (176)$$

where  $\mathbf{e}_k = \mathbf{y}_k - \mathbf{c}(\hat{\mathbf{x}}_{k|k-1})$  is the innovation,  $\mathcal{K}_k = \mathcal{P}_{k|k} \mathcal{C}_k^T \mathcal{R}^{-1}(\mathbf{p}_k)$  is the Kalman gain, and  $\mathcal{C}_k$  is the Jacobian of  $\mathbf{c}(\mathbf{x}_k)$  with respect to  $\mathbf{x}_k$  evaluated at  $\hat{\mathbf{x}}_{k|k-1}$ .

The supervector forms given above generally reduce to independent parallel EKFs for each target provided there is no inter-target coupling. When targets are close, however, this coupling occurs and the supervector form must be maintained in order to account for cross-covariance introduced in the estimation procedure.

#### 4.5. Cluttered Sensing Environment: The MS-JPDAF

In a cluttered sensing environment, tracking algorithms require a data association method to deal with false alarms and missed detections, determine measurement origins, and incorporate these effects into a state estimation procedure. Most of our work has centered on the multisensor joint probabilistic data association (MSJPDA) method first detailed in [64], but our characterization of the problem is compatible with other common methods [65]. More specifically, our simulations employ the sequential MSJPDA [66], which uses multiple single-sensor JPDA methods to sequentially process each sensor's measurements, initializing the next iteration with the filtered outputs of the previous step. This process results in

Assuming an extended Kalman filter (EKF) is employed for the propagation equations, the prediction step is identical to the standard EKF above, resulting in an estimate of the dynamic state  $\hat{\mathbf{x}}_{k|k-1}^i$  and error covariance  $\mathbf{P}_{k|k-1}^i$  for each target  $i$  at time  $k$ . The MSJPDA temporarily treats the various modes of each sensor as independent sensors, resulting in  $N_{\bar{s}} = N_s M$  total sensor-modes. If we define the intermediate filtered estimate and covariance for the  $i$ th target

from the  $j$ th sensor/mode as  $\hat{\mathbf{x}}_{F,j}^i$  and  $\mathbf{P}_{F,j}^i$  respectively (where the subscript  $F$  denotes “filtered”), with initial values  $\hat{\mathbf{x}}_{F,0}^i = \hat{\mathbf{x}}_{k|k-1}^i$ , then  $\mathbf{P}_{F,j}^i = \mathbf{P}_{k|k-1}^i$ . Each update iteration starts by calculating the joint probabilities  $\beta_{l,j}^i$  of the  $l_j$ th measurement being a true measurement of the  $i$ th target (including the no-measurement probability  $\beta_{0,j}^i$ ) as in the standard JPDA [67]. Then, we generate the intermediate updated per-target state estimate and covariance with the recursions

$$\hat{\mathbf{x}}_{F,j}^i = \hat{\mathbf{x}}_{F,j-1}^i + \mathbf{K}_j^i \bar{\mathbf{v}}_j^i \quad (177)$$

$$\begin{aligned} \mathbf{P}_{F,j}^i &= \mathbf{P}_{F,j-1}^i - [1 - \beta_{0,j}^i] \mathbf{K}_j^i \mathbf{S}_j^i \mathbf{K}_j^{iT} \\ &+ \mathbf{K}_j^i \left[ \sum_{l_j=1}^{m_{j,k}} \beta_{l_j,j}^i \mathbf{v}_{l_j,j}^i \mathbf{v}_{l_j,j}^{iT} - \bar{\mathbf{v}}_j^i \bar{\mathbf{v}}_j^{iT} \right] \mathbf{K}_j^{iT} \end{aligned} \quad (178)$$

where we define the per-measurement innovation  $\mathbf{v}_{l_j,j}^i = \mathbf{y}_{l_j,j} - \mathbf{c}_j(\mathbf{x}^i)$  and the combined innovation  $\bar{\mathbf{v}}_j^i = \sum_{l_j=1}^{m_{j,k}} \beta_{l_j,j}^i \mathbf{v}_{l_j,j}^i$ . The filter gain  $\mathbf{K}_j^i$  depends on the underlying estimation procedure, as does the innovation covariance  $\mathbf{S}_j^i$ . Under the EKF assumption,  $\mathbf{K}_j^i = (\mathbf{P}_{F,j-1}^i)^{-1} \mathbf{C}_j^i (\mathbf{S}_j^i)^{-1}$  and  $\mathbf{S}_j^i = \mathbf{C}_j^i \mathbf{P}_{F,j-1}^i \mathbf{C}_j^{iT} + \mathbf{R}_j^i$ , where  $\mathbf{C}_j^i$  is the Jacobian of the measurement function evaluated at  $\hat{\mathbf{x}}_{F,j-1}^i$ . The final updates are given by the filtered values from the last sensor; that is,  $\hat{\mathbf{x}}_{k|k}^i = \hat{\mathbf{x}}_{F,N_S}^i$  and  $\mathbf{P}_{k|k}^i = \mathbf{P}_{F,N_S}^i$ .

#### 4.6. Finite Field-of-View Modeling

The measurement models discussed above implicitly assume that the target states can be observed throughout the entire region of interest, regardless of parameter setting. Furthermore, the noise variance for each measurement,  $\sigma_{meas}^2$ , is presumably uniform throughout the validation region and valid for *all* possible measurement values & sensing parameters. In practice, sensors are limited to a finite field-of-view (FOV) for each measurement they provide, which constrains them to observing only a small subset of the target space. The extent of this FOV is dictated by the same parameters that define the overall measurement variance, thus constraining the feasible set of signal parameters available to the system at each time step.

This fundamental sensing limitation can be handled in a variety of ways. Many algorithms and systems in the existing signal processing literature simply ignore this effect, relying instead on judicious guessing or the presumption of a clairvoyant operator who keeps the sensing parameters within a range so as to maintain a “reasonable” level of coverage of the target space. In [63], Zhang, et al. acknowledge the tradeoff between FOV and measurement accuracy, but do not explicitly model it.

Mathematically, this limitation can be encoded by considering each mode  $j$ ’s FOV as a set  $\mathcal{F}_{j,k}(\mathbf{p}_{j,k})$  that depends on the mode parameters at each time step. Then, the ideal measurement equation of target  $i$  at mode  $j$  can be modified as follows:

$$\mathbf{y}_{j,k}^i = [\mathbf{c}_j(\mathbf{x}_k^i) + \mathbf{v}_{j,k}^i] I_{\mathcal{F}_{j,k}(\mathbf{p}_{j,k})}(\mathbf{x}_k^i) \quad (179)$$

where  $I_{\mathcal{F}_{j,k}(\mathbf{p}_{j,k})}(\mathbf{x}_k^i)$  is the set indicator function given by

$$I_{\mathcal{F}_{j,k}(\mathbf{p}_{j,k})}(\mathbf{x}_k^i) = \begin{cases} 1 & \mathbf{x}_k^i \in \mathcal{F}_{j,k}(\mathbf{p}_{j,k}) \\ 0 & \mathbf{x}_k^i \notin \mathcal{F}_{j,k}(\mathbf{p}_{j,k}) \end{cases} \quad (180)$$

In our framework, we explicitly model this constraint by adding a penalty function  $\phi$  to the measurement variance that assigns large variances to target measurements that would appear outside a particular sensor's FOV. This corresponds to the indicator function-like nature of the constraint: if a target lies within the sensor's FOV, then it is measured with some known variance; those outside the FOV are not measured at all, which contributes to greater uncertainty in the estimate of the state vector. However, since indicator functions are non-differentiable, they are difficult to implement and manage in realistic optimization schemes; hence, we replace the indicator with the penalty function.

More formally, this extended concept of measurement variance can be given as

$$\sigma_{meas}^2 = \sigma_{min}^2(\mathbf{p}) + \phi(x, FOV(\mathbf{p})) \quad (181)$$

where  $x$  is the measured quantity and  $\mathbf{p}$  is, again, a vector of sensor parameters (e.g., pixel angular subtense) that parameterizes both the minimum measurement variance  $\sigma_{min}^2$  and the sensor  $FOV$  (which is a simpler function of the set  $\mathcal{F}_{j,k}(\mathbf{p}_{j,k})$ ). Simple penalty functions include scaled Heaviside step functions (and continuous approximations thereof, which are discussed below), as well as the "deadzone-linear" function [68]. Descriptions of more complicated penalty functions, like the log-barrier and Huber functions, can be found in Boyd & Vanderberghe [68] and references therein. Note that these formulations need to be shifted and/or scaled appropriately for each sensor's particular operation parameters.

In the simulations in Section 4.8., we use a naive continuous approximation to the Heaviside step function, as the discontinuous nature of the step function makes less attractive for optimization. The unit step can be approximated using the hyperbolic tangent function [69]:

$$\begin{aligned} u(x - x_o) &\approx \frac{1}{2} (1 + \tanh(v(x - x_o))) \\ &= (1 + \exp(-2v(x - x_o)))^{-1}. \end{aligned} \quad (182)$$

for large values of  $v$ , a shaping parameter. Examples of the approximation for certain values of  $v$  are shown in Figure 1, which shows clearly that the larger the shaping parameter, the better the approximation. However, this comes with the caveat that a sharper transition near the FOV boundary is less amenable to numerical optimization schemes. With this approximation, our penalty function now becomes:

$$\phi(x, FOV(\mathbf{p})) = \frac{\gamma_x}{1 + \exp(-2\nu(x - FOV(\mathbf{p})))} \quad (183)$$

where, again,  $\gamma_x$  is an appropriately large constant.

Now that we have an appropriate means of modeling the effect of the finite FOV on measurement variance, we can apply it to the RF/EO system whose measurement equations were described in Section 4.5. For the pulse-doppler radar considered, the pulse repetition frequency (PRF) defines the FOV for both the range and range-rate measurements in the form of the maximum unambiguous values  $R_{max}$  and  $\dot{R}_{max}$ :

$$FOV(r) = [0, R_{max}] \quad R_{max} = \frac{\lambda \text{PRF}}{2} \text{ m} \quad (184)$$

$$FOV(\dot{r}) = [0, \dot{R}_{max}] \quad \dot{R}_{max} = \frac{c}{2\text{PRF}} \text{ m/s} \quad (185)$$

where  $\lambda$  is the pulse carrier wavelength and  $c$  is the speed of light. Similarly, for a 1-D EO image that is  $N_{pix}$  pixels wide with a detector (pixel) angular subtense (DAS) of  $\alpha$  radians, the FOV for the azimuth measurement is:

$$FOV(\eta) = \left[ -\frac{N_{pix}\alpha}{2}, \frac{N_{pix}\alpha}{2} \right] \text{ rad.} \quad (186)$$

Since the FOV for the EO sensor is two-sided, the corresponding penalty function will also be "two sided," i.e. the sum of two approximations for each side of the domain.

This penalty formulation forms the backbone of the control algorithm in Section 4.7. and is extensible to other sensing modes, provided there is a generally tractable set description of the FOV. We note that this formulation is reminiscent of "fuzzy set" membership, but we have not explicitly used fuzzy logic/set theory.

#### 4.7. Proposed Closed-Loop Control & Optimization Algorithm

This section describes the proposed algorithm, which incorporates both a covariance control aspect and a parameter optimization aspect to account for the FOV constraint. The context and appropriateness of using this scheme in clutter is also discussed.

##### 4.7.1. Incorporating Finite FOV Effects in State Estimation

The first step of the proposed method incorporates finite field-of-view effects into the tracking scheme by replacing the solely parameter-dependent sensor/mode covariance  $\mathbf{R}_j$  with the parameter and target state dependent  $\tilde{\mathbf{R}}_j^i = \mathbf{R}_j + \Phi_j^i$ , where  $\Phi_j^i$  is a matrix formed from the per-mode penalties above. In the supervector formulation, this equates to replacing the measurement error covariance matrix  $\mathcal{R}$  in the state-space and state estimation/association formulations above with a modified version,  $\tilde{\mathcal{R}}$ , whose elements incorporate the penalty function

as given in Section 4.6. above. Nominally, we would require this true covariance to carry out any optimization or estimation; however, we cannot directly calculate  $\tilde{\mathcal{R}}$  because we do not have access to the target state (and thus the true measurements). Instead, we use an estimate,  $\hat{\mathcal{R}}$ , that is obtained by evaluating the penalty functions at the predicted next-state measurements  $\hat{\mathcal{Y}}_{k+1|k}$ . This estimated covariance is then propagated through both the state estimation procedure and the optimization scheme discussed below.

#### 4.7.2. Adaptive Parameter Optimization Scheme

As with any tracking system, our purpose in developing this control algorithm is to somehow improve tracking performance, which is usually characterized in perfect detection environments by minimizing mean-square tracking error. Likewise, our method selects the next-step sensing parameter vector  $\mathbf{p}_{k+1}$  (and thus, indirectly, each mode's FOV) that minimizes, in a sense, an equivalent metric: a scalarization of the next-step filtered covariance matrix  $\mathcal{P}_{k+1|k+1}$ . One commonly proposed scalarization is the trace weighted by an appropriate positive semidefinite matrix  $\mathbf{W}$  [53], [70], which assigns a designer-imposed importance value to particular components of the state vector or targets. An identical method, instead of minimizing the next-step prediction covariance, would be to *maximize* some scalarization of the predicted *information* matrix  $\mathcal{P}_{k+1|k+1}^{-1}$ , or a component thereof [70]. We consider modifications of this cost function below. This method is related to other recent advances in sensor management (i.e. Kreucher, *et al.* [71]), where the objective function derives from more complex information metrics like the Renyi divergence. In the case of the linear/non-linear Gaussian assumption, many of these metrics collapse to the predicted information matrix.

##### 4.7.2.1. Main Objective Function Selection

The general form of the optimization problem we solve depends on the tracking matrix considered. Using the next-step filtered covariance matrix  $\mathcal{P}_{k+1|k+1}$  implies we consider a minimization of the scalarization function. Conversely, if we scalarize the filtered information matrix  $\mathcal{P}_{k+1|k+1}^{-1}$ , then the goal is to maximize said scalarization. This latter form is the basis for the objective functions considered below.

The first objective function is derived using equivalent maximizations and some approximations. If we initially choose the determinant as our scalarization, so the objective function is  $|\mathcal{P}_{k+1|k+1}^{-1}|$ , we note that maximizing this quantity w.r.t  $\mathbf{p}_{k+1}$  is equivalent to maximizing  $\log|\mathcal{P}_{k+1|k+1}^{-1}|$ . It can be shown that, using a Taylor series approximation and assuming the spectral radius of the information increase matrix  $\mathcal{C}_{k+1}^T \hat{\mathcal{R}}^{-1}(\mathbf{p}_{k+1}) \mathcal{C}_{k+1}$  is sufficiently small, maximizing the log determinant is equivalent to maximizing  $f_{0,Appx}(\mathbf{p}_{k+1}) = \text{tr}(\mathcal{P}_{k+1|k} \mathcal{C}_{k+1}^T \hat{\mathcal{R}}^{-1}(\mathbf{p}_{k+1}) \mathcal{C}_{k+1})$ . We term this the "approximation" objective function. This cost function is also reminiscent of the aforementioned weighted trace of the filtered covariance; here, however, we weight the next-step *information increase* with the *prediction error*.

We obtain the second objective function in our analysis in a much more straightforward manner. Following the concepts demonstrated in [70], [72], we note that the only portion of the perfect-



detection filtered multitarget information matrix  $\mathcal{P}_{k+1|k+1}^{-1} = \mathcal{P}_{k+1|k}^{-1} + \mathcal{C}_{k+1}^T \hat{\mathcal{R}}^{-1}(\mathbf{p}_{k+1}) \mathcal{C}_{k+1}$  that depends on the parameter vector is the right hand term, the so called “information increase” after the filtering step of the Kalman filter. Hence, instead of maximizing a scalarization of the entire information matrix, we only seek to maximize a scalarization of the information increase, hereafter referred to as MII. The associated scalarization chosen here is the determinant, thus the given objective function is  $f_{0,MI}(\mathbf{p}_{k+1}) = |\mathcal{C}_{k+1}^T \hat{\mathcal{R}}^{-1}(\mathbf{p}_{k+1}) \mathcal{C}_{k+1}|$ . Note that this objective does not include the prediction error which, as we will see later, affects the convergence of the algorithm.

In either case, the basic optimization problem is now:

$$\begin{aligned} \max_{\mathbf{p}_{k+1}} \quad & f_0(\mathbf{p}_{k+1}) \\ \text{s. t.} \quad & \mathbf{l} \preceq \mathbf{p}_{k+1} \preceq \mathbf{u} \end{aligned} \tag{187}$$

where  $f_0(\mathbf{p}_{k+1})$  is one of the objectives described above, the vectors  $\mathbf{u}$  and  $\mathbf{l}$  contain the maximum & minimum possible sensor parameter values, respectively, and the operator  $\preceq$  denotes an elementwise vector inequality.

#### 4.7.2.2. Minimum Information Quality: Soft Constraints

While we readily admit that the fundamental conception of the initial optimization problem is seemingly basic, the field-of-view limitation adds a significant wrinkle to the situation heretofore unexamined. This problem is that, even with the penalty function formulation, merely maximizing the information metrics in the previous subsection could potentially drive multiple sensor/modes to low variance/highly restricted FOV settings, focused on a single target (or indeed, *no* target if the estimation procedure diverges) at the expense of the remaining targets to be tracked. That is, it is possible for  $\mathbf{p}_{k+1}$  to maximize the overall cost function at the expense of providing almost no new information about a particular target. This leads to an overall performance degradation that may not be accurately reflected in the primary objective function. This is clearly behavior we wish to avoid, so it becomes necessary to introduce an additional constraint

However, one should not take this to mean the solution is to artificially constrain every mode to maintain every target within its view at all times. Indeed, such a constraint would defeat the purpose of the penalty functions, whose intent is to encode the indicator non-linearity in a “soft” manner. Instead we recognize that it is more important to maintain each target’s overall track quality, and let that be the guiding principle behind the constraint.

Given the objective functions, we can derive nonlinear constraints for the optimization that avoid mode hyperfocus. Noting that the matrices are all block diagonal, partitioned by the associated target, we observe that each objective function is composed of a sum or product of similarly constructed per-target functions. Since we wish to maintain a certain baseline amount of information, we impose a lower threshold  $t$  on the smallest per-target function, corresponding to the target that contributes the least in the next step. Assuming that the function  $g^{(i)}(\cdot)$  is

functionally identical to the primary objective  $f_0(\cdot)$ , but operates on the per-target matrices of the  $i$  th target, we can express this new constraint as

$$\min_i g^{(i)}(\mathbf{p}_{k+1}) \geq t \quad (188)$$

With this new constraint, the final optimization problem becomes:

$$\begin{aligned} & \max_{\mathbf{p}_{k+1}} f_0(\mathbf{p}_{k+1}) \\ \text{s. t. } & \min_i g^{(i)}(\mathbf{p}_{k+1}) \geq t \\ & \mathbf{l} \preceq \mathbf{p}_{k+1} \preceq \mathbf{u}. \end{aligned} \quad (189)$$

#### 4.7.3. Algorithm Modifications for Cluttered Sensing

As in the ideal sensing case, tracking metrics like mean-square state estimation error and next-step information increase [72] remain important, since accurately determining target position is a primary purpose of these types of data fusion systems. However, in uncertain measurement environments, it is also critical to examine costs that include association metrics. An example objective might be to maximize the overall probability of correct association for a particular (sub)set of sensing modes. For soft association methods like the one presented above, however, analytically formulating this metric is difficult due to the non-deterministic nature of measurement-to-track hypothesis generation. Furthermore, it is not clear that improving association without regard to the tracking performance has many, if any, benefit. A hybrid approach to managing the tracking and association goals is to optimize a function of the validation regions for each mode, as advanced in [53], which affects both association probabilities and tracking error.

With these concepts in mind, we instead chose to naively apply the same algorithm to the cluttered case, with minor modifications for the particularities of the state estimation. The covariance control aspect augments the MSJPDA/EKF scheme from Section 4.5 by replacing the solely parameter-dependent sensor/mode covariance  $\mathbf{R}_j$  with the parameter and target state dependent  $\tilde{\mathbf{R}}_j^i = \mathbf{R}_j + \Phi_j^i$ , where  $\Phi_j^i$  is a matrix formed from the per-mode penalties above. However, since  $\tilde{\mathbf{R}}_j^i$  requires evaluating the penalty at the actual target state (which is unknown), we instead use the estimates of each target state from each stage of the sequential MSJPDA, and thus introduce an estimated penalized covariance  $\hat{\tilde{\mathbf{R}}}_j^i$ . This has the effect of reducing the validation region size as each subsequent mode's measurements are processed.

The parameter optimization problem remains essentially unchanged. Incorporating the constraints, the final optimization problem is as in Section 4.7.2:

$$\begin{aligned} & \max_{\mathbf{p}_{k+1}} f_0(\mathbf{p}_{k+1}) \\ \text{s. t. } & \min_i g^{(i)}(\mathbf{p}_{k+1}) \geq t \\ & \mathbf{l} \preceq \mathbf{p}_{k+1} \preceq \mathbf{u}. \end{aligned} \quad (190)$$

where  $\preceq$  indicates elementwise inequality,  $t$  is the minimum information threshold, and  $\mathbf{l}, \mathbf{u}$  are system-imposed lower & upper bounds on the parameter vector. The block matrices  $\mathcal{P}_{k+1|k}$ ,  $\mathcal{C}_{k+1}^T$  and  $\hat{\mathcal{R}}^{-1}(\mathbf{p}_{k+1})$  that appear in the two cost functions possible for  $f_0(\mathbf{p}_{k+1})$  are formed by concatenating first by sensor, then by target, the per-target/per-sensor error covariances  $\mathbf{P}_j^i$ , measurement Jacobian matrices  $\mathbf{C}_j^i$ , and augmented measurement covariances  $\hat{\mathbf{R}}_j^i$  (evaluated at the next-step predicted state  $\hat{\mathbf{x}}_{k+1|k}^i$ ), respectively. All of these values are derived from their final estimates at their stage in the MSJPDA process. It may seem short-sighted to employ the same optimization problem for the cluttered scenario, since the above described optimization problem was initially constructed to minimize tracking error in ideal detection/association situations, but as we will show in Section 4.8., these cost functions remain effective in clutter and result in benefits for association as well.

We note that it is possible to augment our cost functions with the hybrid covariance approximations or modified Riccati equations presented in [66], but these result in significantly more complex cost functions that require further investigation. Preliminary investigations demonstrated that incorporating the approximation is necessary in both the covariance update and the optimization cost function; otherwise, the cost function under-estimates the resulting covariance decrease and provides inconsistent results.

## 4.8. Simulations and Results

This section describes a variety of comparative analyses for the algorithm developed during the project. These analyses demonstrate the utility of the algorithm described above, as well as the necessity of incorporating field-of-view effects into multimodal sensing systems.

Many of the simulations below share common parameters and values, regardless of the target scenario or sensing environment considered. Unless otherwise specified, the simulations consider a tracking period of 500 seconds with a sample period of one second. Two RF/EO multi-mode sensors -- located at the  $(x, y)$  positions  $(-1 \text{ km}, -1 \text{ km})$  (Sensor 1) and  $(1 \text{ km}, -1 \text{ km})$  (Sensor 2) -- are stationary with X-band ( $f_c = 10 \text{ GHz}$ ) radars with pulse-doppler processing & one-dimensional 1000 pixel EO modes with a detector-angular-subtense (DAS) between  $10^{-5}$  and  $3 \times 10^{-3}$  radians/pixel and a look direction of 0 degrees. At each time step, the RF sensors provide range and range-rate measurements, while the EO sensors provide azimuth measurements. The sensor parameters are initialized in varying configurations, but generally in a manner that ensures the targets are initially contained within their FOV.

### 4.8.1. Ideal Sensing Environment Analysis

This subsection covers the primary tracking results for the ideal measurement scenarios. The general simulation parameters described above apply unless otherwise noted. For comparison, we also commonly simulate the scenario and generated target tracks for a similar non-adaptive method.

#### 4.8.1.1. Primary Multiple Target Tracking Results

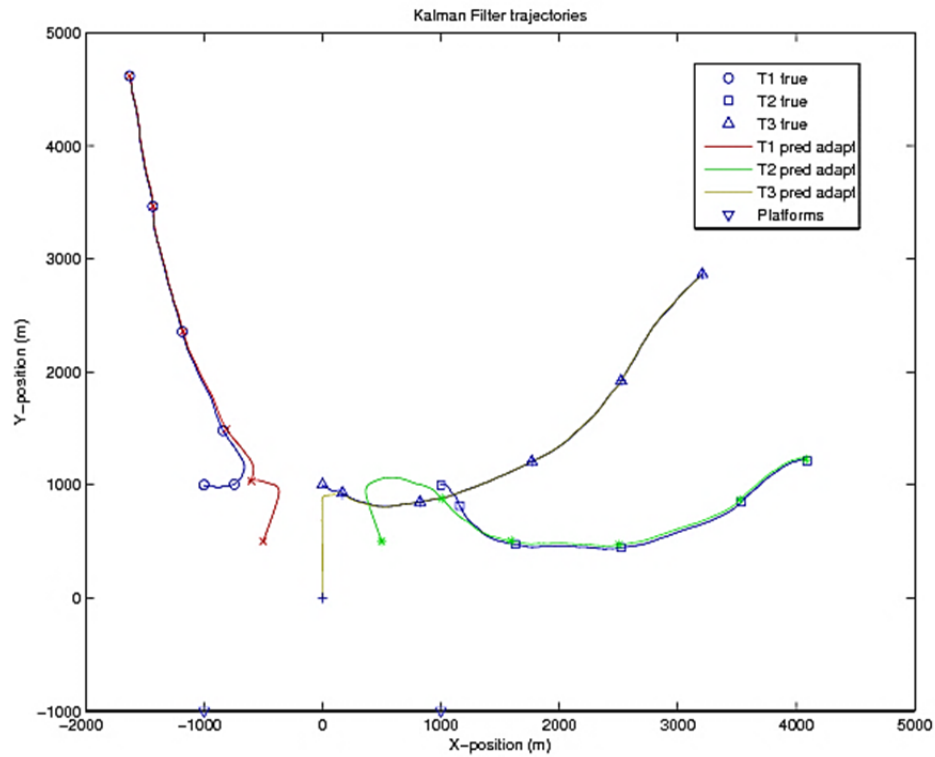
The results considered here demonstrate the tracking performance of the proposed method in an ideal sensing scenario with three targets in the region of interest.

Beyond the common simulation parameters, this example assumes the following: Both sensors were initialized with identical system parameters (PRF and bandwidth for the RF sensor, DAS for the EO sensor). The starting positions for the targets are at (-1 km, 1 km) (Target 1), (1 km, 1 km) (Target 2), and (0 km, 1 km) (Target 3), which then move with randomly generated trajectories that evolve according to the constant-velocity state model with a multiplicative variance  $q = 0.1$ . Initial position estimates were chosen to be about midway to the actual positions with a large error variance. For comparison, we also simulated the scenario and generated target tracks for a non-adaptive method; that is, where the sensor parameters remained constant.

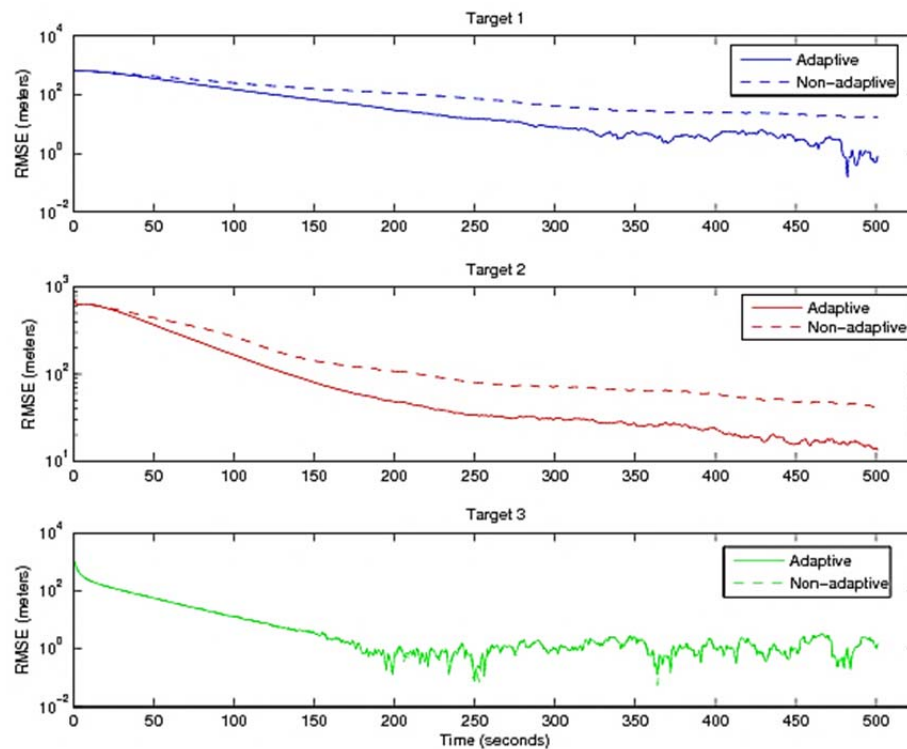
Figure 10 provides a visualization of the tracking performance of the adaptive method, showing the actual and predicted target positions over the entire simulation, as well as the locations of the sensors. For clarity, the actual (or predicted) position of the target is indicated with a marker that is either indicated in the legend or in the color of the track plot. Clearly, the adaptive method is able to handle a situation where the number of targets exceeds the number of sensors, as all the trajectories of all three targets are able to be followed by the adaptive method.

Figure 11 summarizes the RMS predicted position error for each target throughout the simulation for both the adaptive and non-adaptive methods, numerically confirming the results from Figure 10. Both targets 1 and 2 show noticeable improvement in tracking error with the adaptive method over that of the non-adaptive method while target 3's tracking error remains quite low. Due to the initial sensor parameters given and the movement of target 3, the difference is not as noticeable.

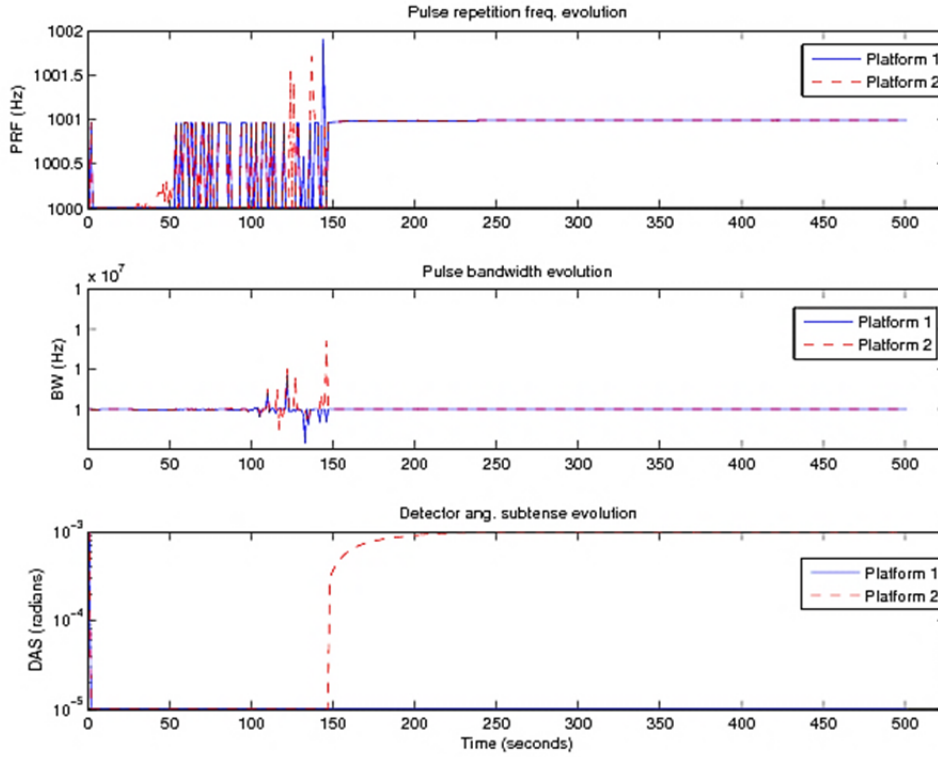
Since the tracking performance seems accurate, it is critical to analyze *how* the sensors measured these tracks, and thus, the obtained policies. Figure 12 displays the evolution of the sensor parameters for the adaptive method throughout the tracking period. As the ranges of all 3 targets increases relative to the sensors (particularly from 50-150 seconds), we can see that the algorithm once again follows the known ranging technique of alternating PRFs until a certain quiescent point is reached (in terms of per-step error/information). Since the initial bandwidth was relatively high (10 MHz, which corresponds to about 19 meters in range variance), there was little need for variation, except towards the end of the ranging period where more accurate range information is required. Finally, we can see that evolution of each EO sensor's DAS demonstrates a tradeoff between sensors. Since Target 1's motion is primarily directly above Sensor 1, that sensor narrows its field of view (but also obtains finer azimuth measurements) by lowering the DAS. This is also initially the case for Sensor 2, as Targets 2 and 3 have restricted angular movement. Past a certain point (~150 seconds), however, these targets rapidly move out of this FOV, which prompts the algorithm to select a larger DAS (and thus a wider FOV); this continues until the maximum value is reached as the azimuth of these targets increases relative to Sensor 2.



**Figure 10. Three-Target Ideal Sensing Scenario Tracks**



**Figure 11. RMS Position Error for Three-Target Ideal Scenario**



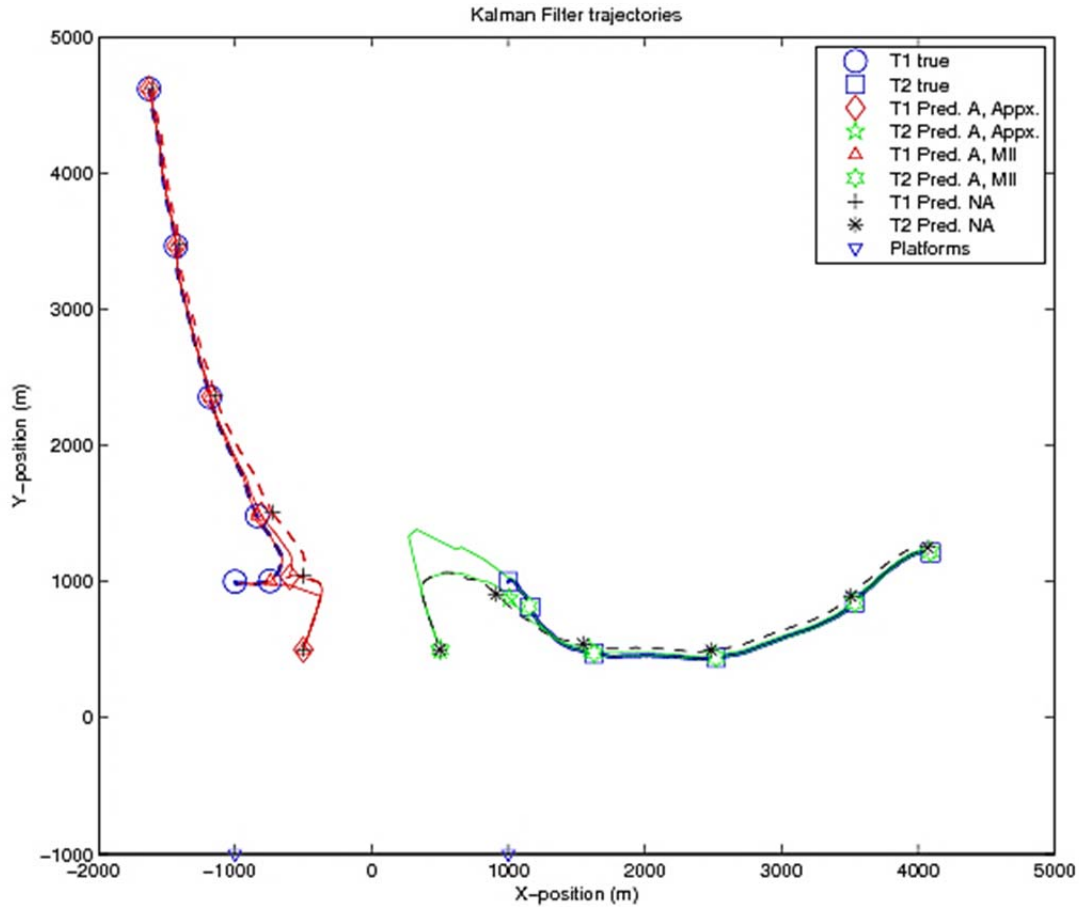
**Figure 12. Parameter Evolution for Three-Target Ideal Scenario**

#### 4.8.1.2. Comparison of Objective Functions

This analysis demonstrates the comparative performance in the ideal sensing scenario between the two objective functions (trace approximation and MII) discussed in Section 4.7. In this case, we analyze a two target scenario consisting of targets 1 and 2 from the analysis above.

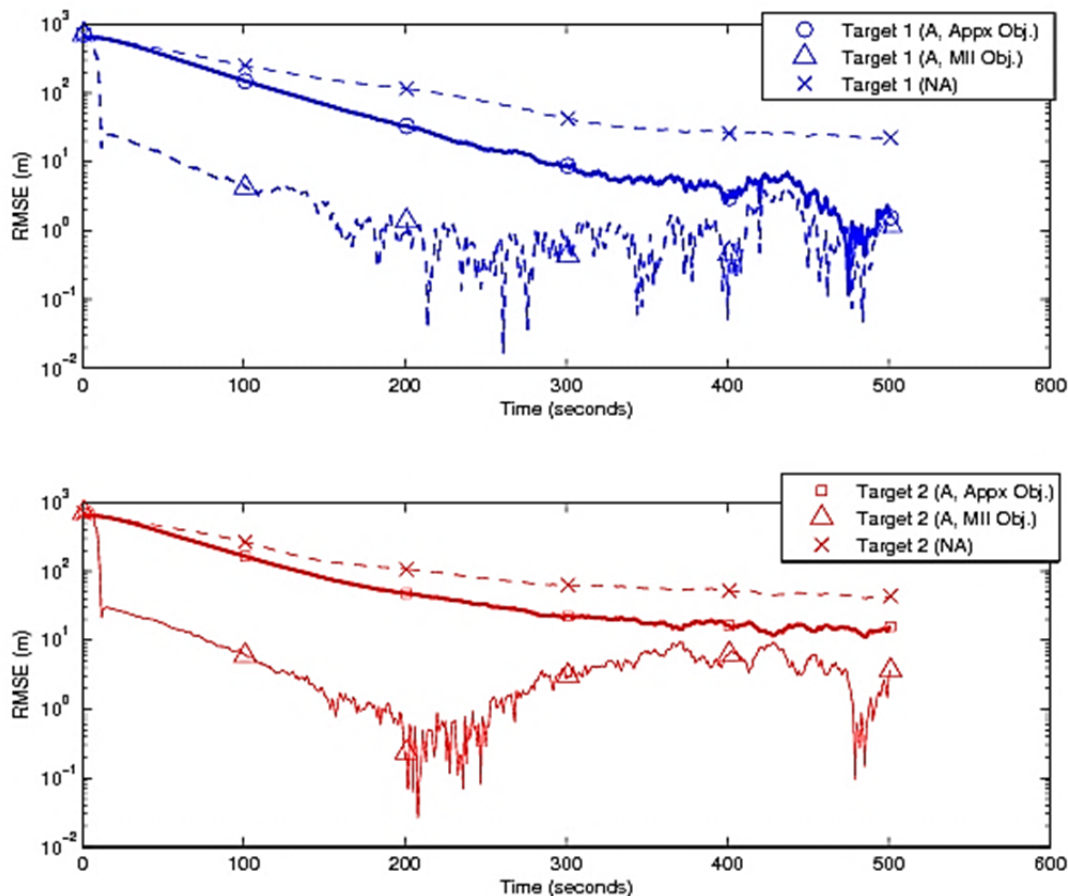
In addition to the common simulation parameters, these results assume the following: The starting positions for the targets are at (-1 km, 1 km) (Target 1), (1 km, 1 km) (Target 2), and (0 km, 1 km) (Target 3), which then move with randomly generated trajectories that evolve according to the constant-velocity state model with multiplicative variance  $q = 0.1$ .

Figure 13 visualizes the tracking performance of the two adaptive, as well as the non-adaptive method. Each track is denoted by a path, supplemented by a marker in the legend denoting the method employed. In the legend, “A” denotes the tracks formed from the two adaptive methods, using either the approximation objective (“Appx.”) or the maximum information increase objective (“MII”). The sensor locations are also marked as blue triangles for comparison. For both targets, we can see that the adaptive methods both converge much quicker than the non-adaptive method on the actual target track, and follow it much more closely. Comparing the adaptive methods, we see that using the MII objective function results in faster track convergence.



**Figure 13. Target Tracks for Objective Function Comparison Analysis**

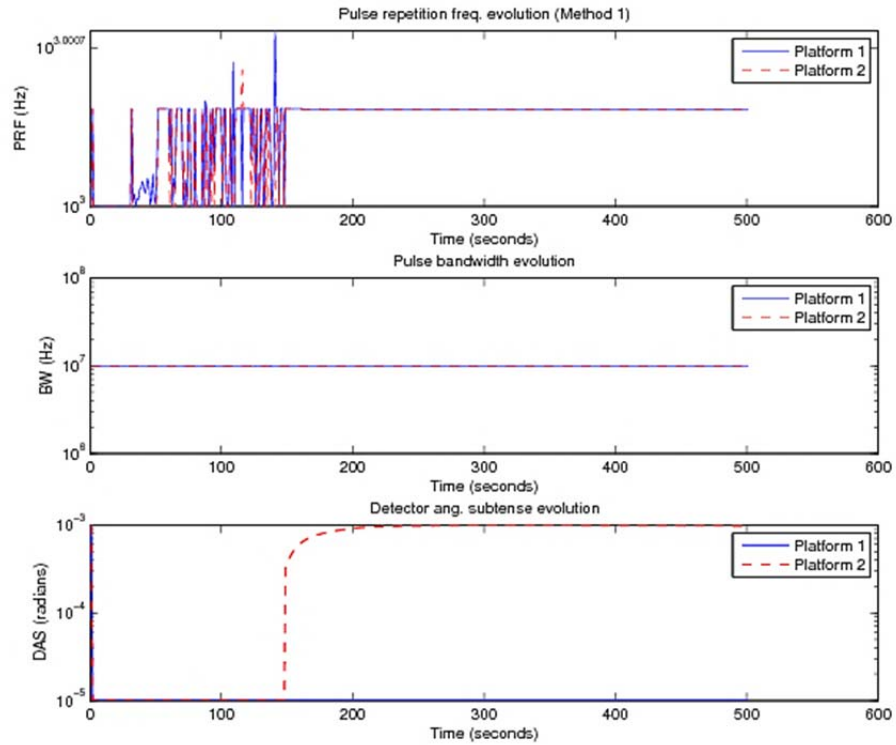
Figure 14 summarizes the RMS predicted position error for each target throughout the simulation for the adaptive and non-adaptive methods, numerically confirming the results from Figure 10. It is clear that adaptation using the MII objective function results in much faster convergence than either the non-adaptive method or the adaptive method with the approximation objective. Furthermore, using the MII function results in lower positioning error than the other methods despite a period of apparent divergence around 200 seconds.



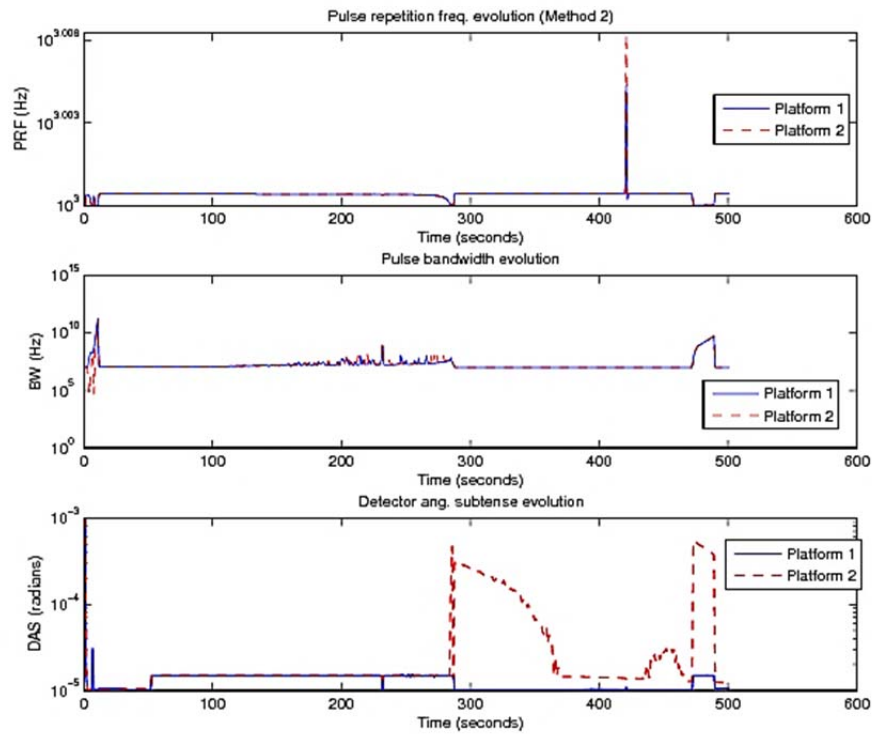
**Figure 14. RMS Position Errors for Objective Function Comparison**

In order to complete the comparison, it is again critical to analyze the obtained sensing policies. Figure 15 and Figure Figure 16 display the sensor parameters for the adaptive methods throughout the tracking period corresponding to the approximation and MII objective functions, respectively. The parameter evolution for the first objective adaptive method is nearly identical to the previously investigated situation in the previous subsection, with classical rangefinding behavior in the pulse repetition frequency (PRF), relatively static pulse bandwidth, and one EO mode (in this case, Sensor 2's) gradually widening its FOV to capture a laterally moving target (here, Target 2). The picture presented for the MII objective function, however, is less intuitive. Some rangefinding behavior in the PRF occurs at the beginning, with small follow-ups throughout the tracking period. This limited initial rangefinding is counterbalanced by a ramping up of pulse bandwidth in the beginning period; later, the bandwidth decreases to close to the original setting, with some perturbations. In terms of the EO modes, the MII optimization scheme chooses to widen *both* sensors' optical FOV slightly between 50 to 275 seconds, corresponding to both targets having significant lateral movement. As Target 1's motion becomes more horizontal (and thus more "in front" of Sensor 1), Sensor 1's EO mode focuses in on the target; conversely, Sensor 2 broadens its FOV further to manage the continued lateral movement of Target 2.





**Figure 15. Parameter Evolution for Adaptive Method with Approximation Objective**



**Figure 16. Parameter Evolution for Adaptive Method with MII Objective**

From these simulations, it is clear that the MII approach provides better tracking performance overall; however, this is but one possible example of the benefit of appropriate selection of objective function.

#### 4.8.2. Cluttered Sensing Environment Analysis

As in the perfect sensing examples above, the simulations considered below share the common simulation parameters described at the beginning of the section. There are also simulation parameters common to the cluttered environment analysis. In most cases, the probability of detection is  $P_D = 0.9$  for all sensor/modes. In addition, the clutter densities (within the sensor's FOV) for the RF and EO sensing modes are  $\bar{\lambda}_{RF} = 2 \times 10^{-7}$  and  $\bar{\lambda}_{EO} = 10^{-5}$ , respectively.

##### 4.8.2.1. Multiple Target Tracking Results

This subsection presents the general tracking results for our proposed algorithm in a cluttered sensing environment. The target scenario considered is the familiar two target randomly-generated trajectory example of the ideal-sensing sections. All parameters and starting points are as detailed above.

Figure 17 through Figure 23 detail the target trajectories, tracking error, and parameter evolution of an example run of the tracking algorithm employing the MII objective criterion described in Section 4.7. The reason for this objective will be made clear later in the section. As can be seen in Figure 17, the non-adaptive method has significant problems converging to the correct target trajectories, experiencing track loss very quickly as the targets move out of the initial FOV and additional false measurements are obtained. Comparatively, the adaptive method's tracks, as shown in Figure 18, converge quickly and accurately to the true trajectories. Confirming this observation, both the position (Figure 19) and velocity (Figure 20) errors are quite significant for the non-adaptive method, but rather manageable (within a meter or so) for the adaptive scheme. Figure 21 through Figure 23 provide us with a look at the evolution of the sensor parameters. For the RF sensor, we can clearly see constant ranging behavior (alternating between higher and lower PRFs), as well as an initial trade-off period of pulse bandwidths (and thus range resolutions). For the EO sensor, we can see that the system opens the FOV by increasing the DAS at points where the target motions deviate from the assumed linear dynamics.

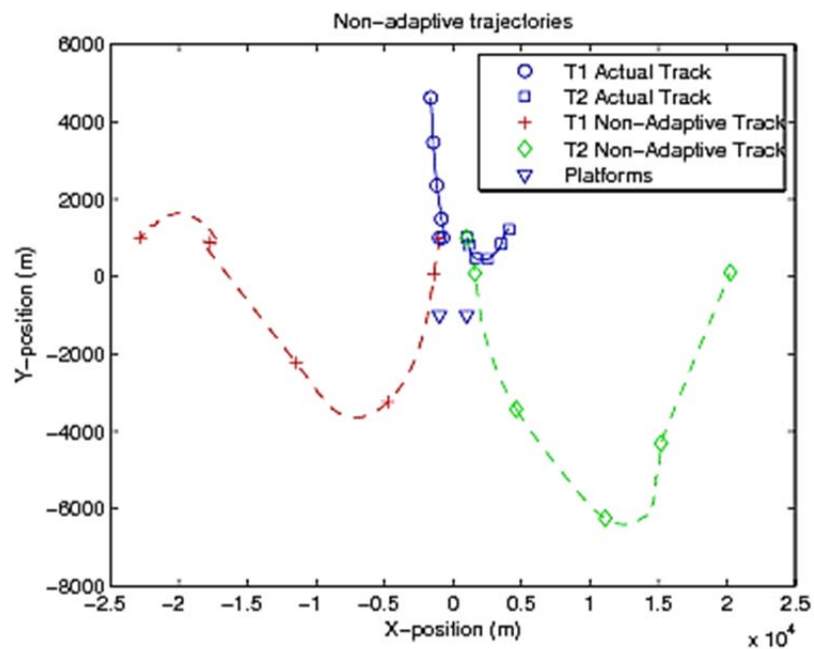


Figure 17. Target Tracks for Non-adaptive Method in Clutter

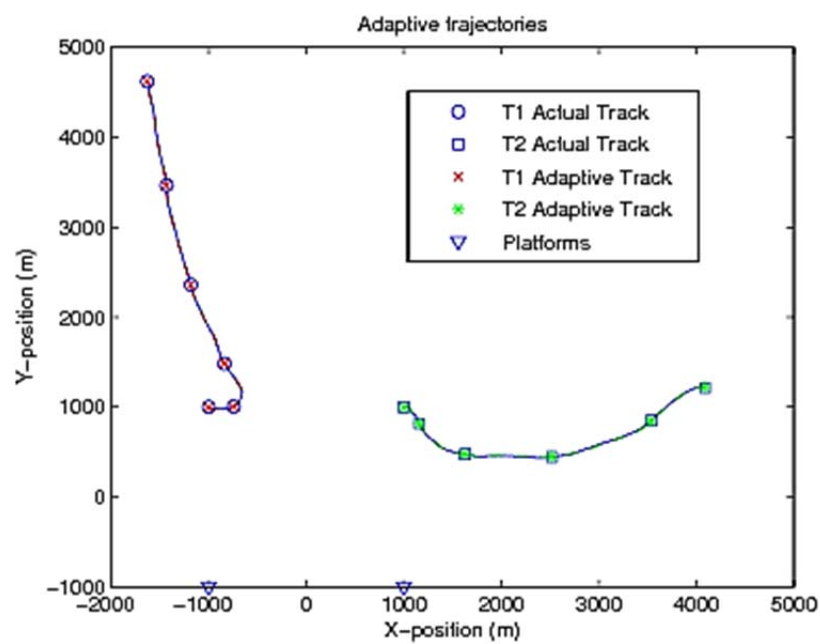
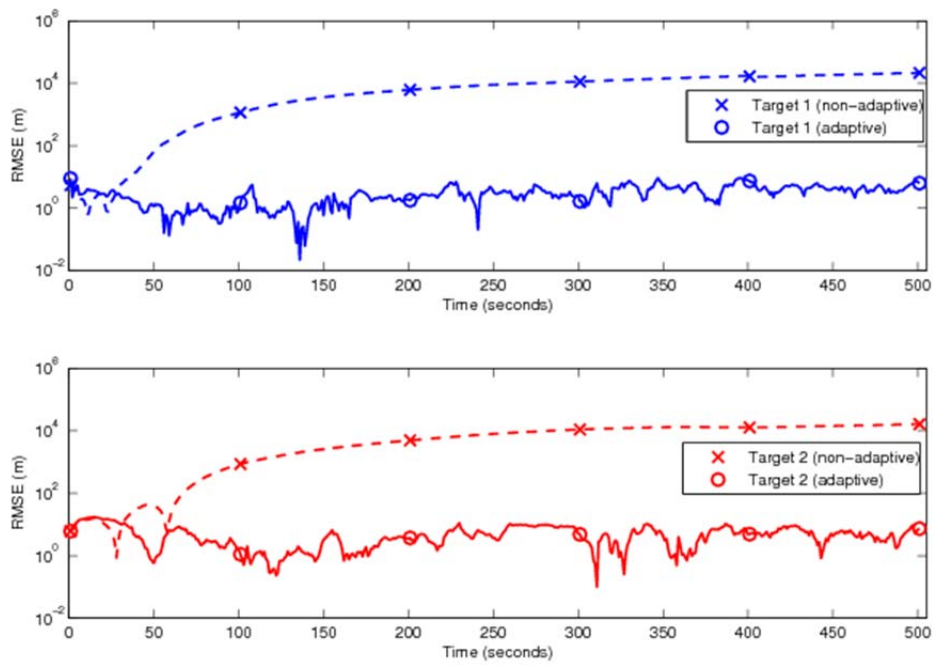
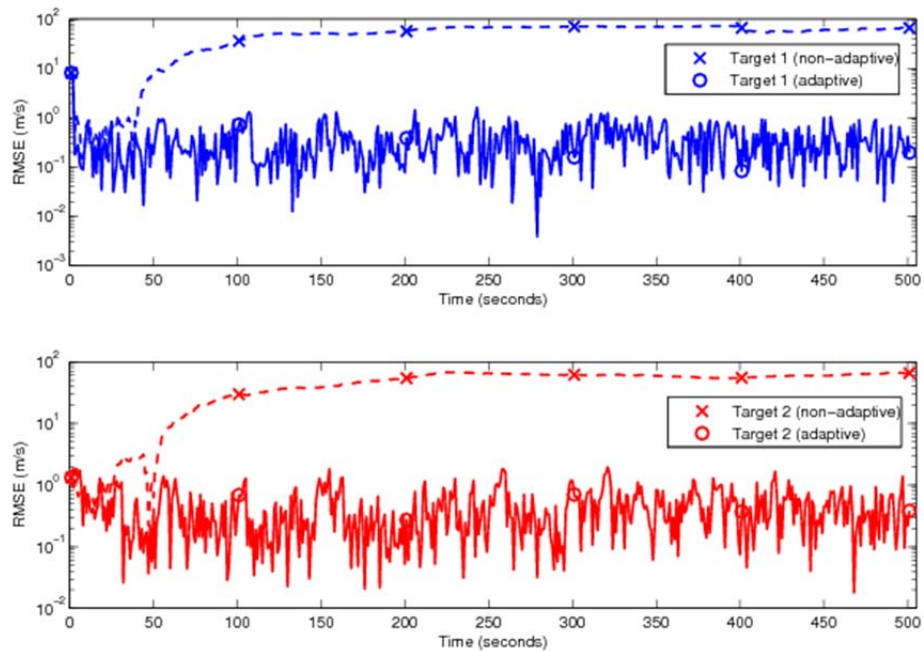


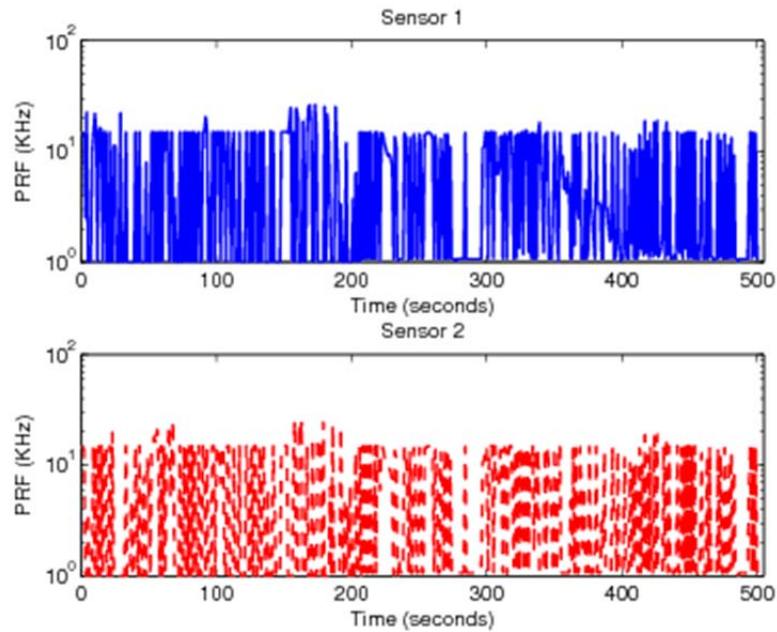
Figure 18. Target Tracks for Adaptive Method in Clutter



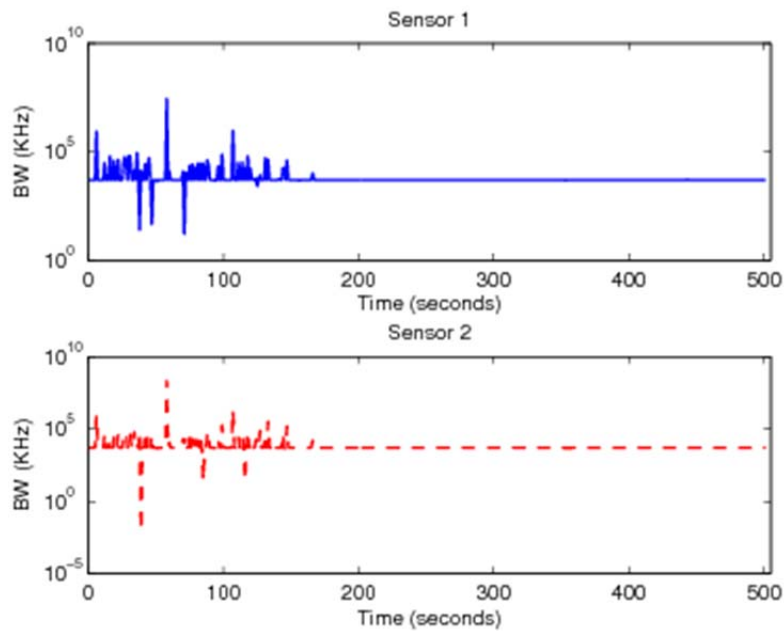
**Figure 19. RMS Position Error for Clutter Scenario**



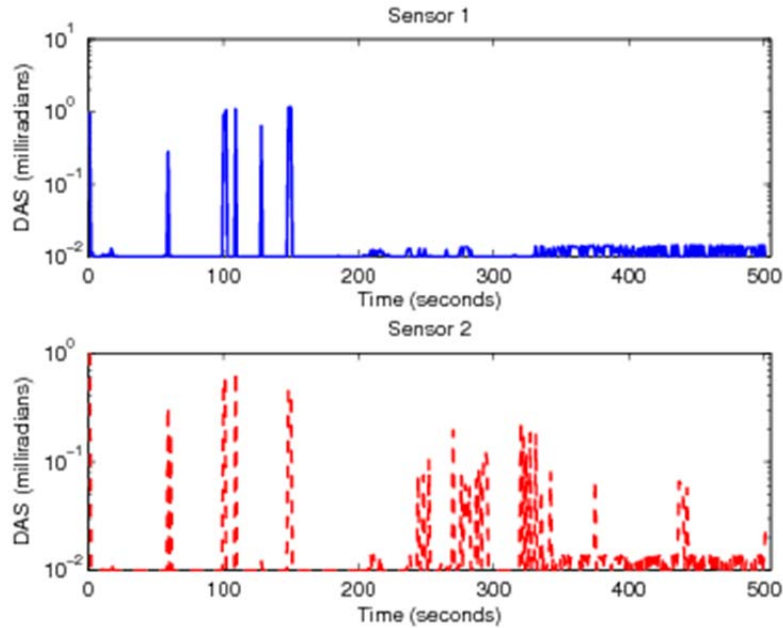
**Figure 20. RMS Velocity Error for Clutter Scenario**



**Figure 21. Adaptive PRF Evolution for RF Mode in Clutter**



**Figure 22. Adaptive Bandwidth Evolution for RF Mode in Clutter**



**Figure 23. Adaptive DAS Evolution for EO Mode in Clutter**

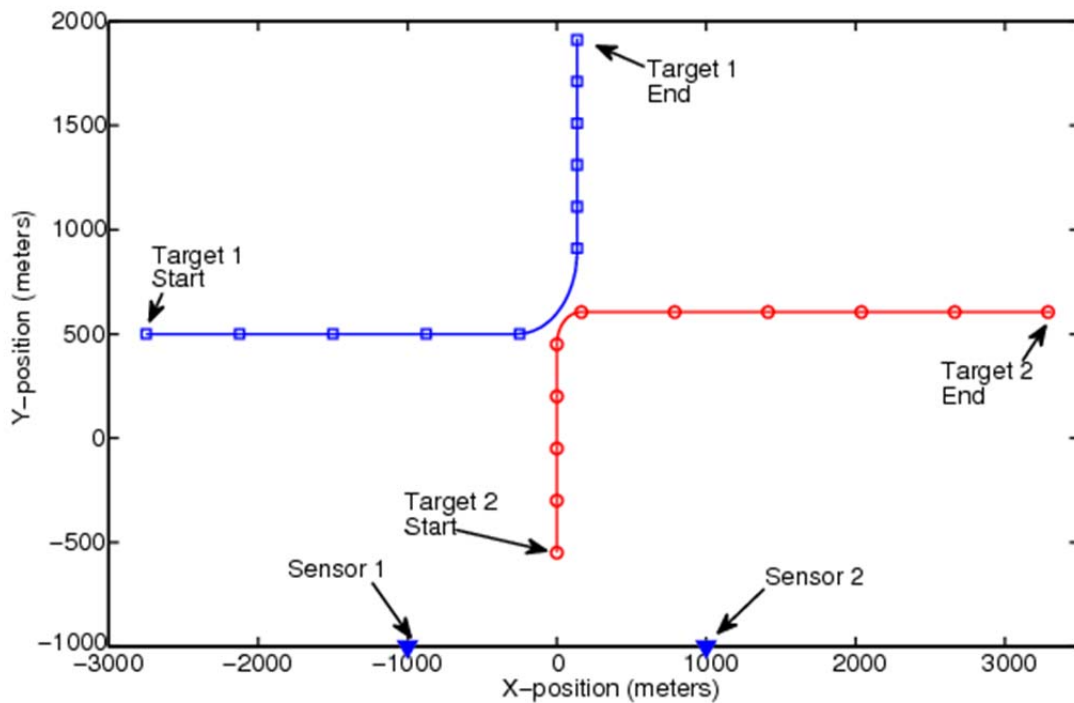
We note that this general behavior holds over many Monte Carlo trials regardless of the objective, but with a few caveats. First, the worst-case target constraint threshold must be set smaller than in the ideal sensing environment as the initial uncertainty about the target state increases to maintain stability. This is due to the normalizing effect of the error covariance on the constraint objective. Second, as with the non-cluttered scenario, the initial set of sensor parameters (as well as where the optimization point method is initialized, though these are usually the same) must be chosen carefully so as to provide as high-quality initial measurements as possible and to avoid local minima in the optimization method. Finally, we note that FOV adaptation using the determinant approximation objective has the tendency to favor one target over the other, which can lead to divergence for a subset of the targets considered. This is likely due to the approximation breaking down in the presence of the clutter and the decreased information gains imposed by the measurement uncertainty.

#### **4.8.2.2. Improvements in Data Association**

In this section, we consider the potential benefit of our method to data association as compared to a non-adaptive method. In previous technical reports and papers, the primary goal has been to track the dynamic states of a set of targets with the least amount of error possible. While this is indeed a reasonable goal, and one that we have generally achieved with our adaptive method, there are other metrics by which we can judge the success/benefit of a given tracking algorithm. This is especially true if we analyze cluttered measurement environments, where the importance of correct data association cannot be discounted. While data association does have a major impact on tracking error, etc., it also sheds light on additional effects of adaptation that may not be examined through tracking error alone, like the impact of sensor field-of-view on measurement validation and clutter reduction.

There are many methods to evaluate & compare the data association capabilities of various tracking algorithm, but a simple first-pass metric is the single-scan misassociation probability (and related metrics of correct and no association probabilities). As the name implies, this measures the probability of incorrectly associating a measurement to a target that did not originate it. We can surmise that if one method provides lower misassociation (higher correct association) probabilities than another, then the first method is superior to the second in some fashion. We note that our current method of data association, in regards to direct measurement-to-target association, appears difficult to directly analyze with this metric, primarily because JPDA is a soft association method. That is, it does not directly provide measurement-to-target associations; rather, it calculates association probabilities and then uses these to weight each measurement's contribution to the state update. Despite this apparent contradiction, we can still create equivalent hard associations by assigning measurements to their most probable source. In practice, this is accomplished by determining the hypothesis with the maximum association probability for each measurement, and declaring that as the obtained association.

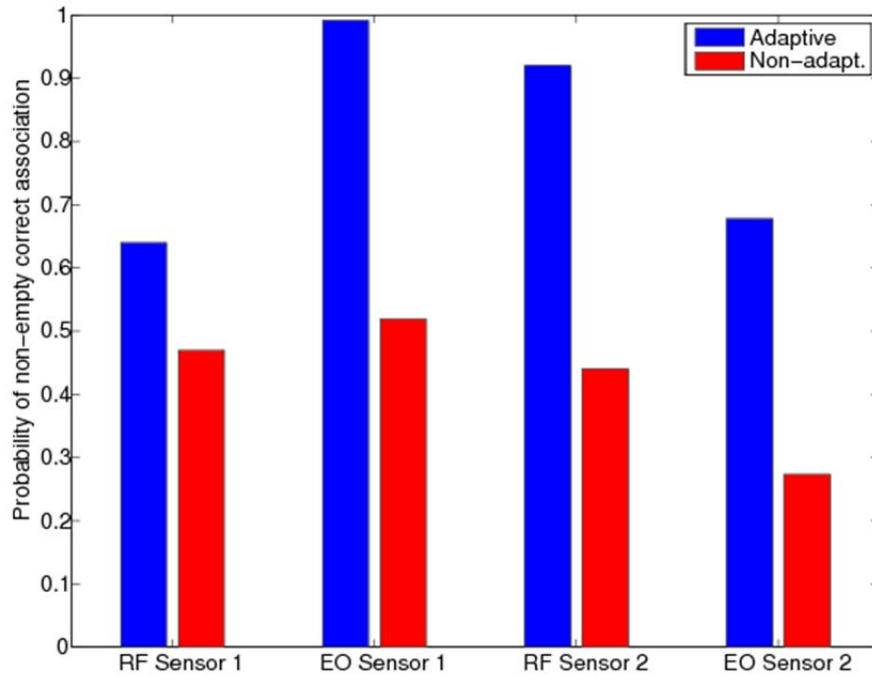
For this simulation and all subsequent simulations, we consider a novel maneuvering target scenario: two targets that nearly cross, so as to provide a major association & validation challenge for our method. The scenario is illustrated in Figure 24 with Target 1 in blue, Target 2 in red, and their respective markers indicating position every 50 seconds, starting from the left (bottom) for Target 1 (2). Target 1 moves rightwards at 12.5 m/s from 0 to 200s, executes a 1.8 degree/sec counterclockwise turn for 50s, and then moves upwards at 4 m/s until the end of the tracking period. Target 2 moves upwards at 5 m/s from 0 to 200s, then executing a 1.8 degree/sec clockwise turn for 50s, and finally moves rightwards at 4 m/s until the end of the tracking period. During their coordinated turns, the targets are about 50-100 meters apart. In this case, we consider the approximation objective which, as we will see, handles this scenario in a slightly more graceful manner than the randomly-generated trajectory instance.



**Figure 24. Maneuvering Target Scenario Trajectories for Clutter Analysis**

Figure 25 below demonstrates the association results from an example 200-iteration Monte Carlo run, plotting each sensor type's average probability of correct association for the adaptive method under consideration and a non-adaptive method. It is clear that for every sensor considered, the data association is greatly improved by the adaptive method. Note, however, that this quantity is specified for "non-empty" associations; that is, periods that had measurements. This is a particularly important point, because the adaptation can result in an extended period of no measurements for a given sensor. This occurs when the chosen field-of-view is so small that the targets are out of view and the expected number of false measurements is nearly zero. In our analysis, it appears that extended no-measurement (and thus no-association) periods are quite common for EO sensors in the adaptive scheme, with no-measurement periods extending to 90% of the tracking period in some cases. For the non-adaptive scheme, extended no-measurement periods (again, particularly focusing on the EO sensors) are comparatively rare, occurring about 20% of the tracking periods.





**Figure 25. Probability of Non-Empty Association Comparisons**

This behavior may seem strange, but in terms of our algorithm, there is a basis for it: Since the goal of the objective function is to maximize the information increase subject to a *minimum per-target threshold*, the adaptive scheme instead shuts out sensors that cannot meet the threshold (and thus contribute to accurate target measurements) to avoid introducing further clutter & error into the state estimation. By comparison (and this is verified by our simulations), the non-adaptive method suffers by keeping the EO modes' FOV large, leading to misassociation with clutter and incorrect target associations, especially during the turning maneuvers. This also applies to the RF modes to a lesser extent, particularly in the range measurement.

#### 4.8.2.3. Necessity of Adaptivity and Compensation: Algorithm Comparisons

This subsection demonstrates the necessity of both aspects of our proposed algorithm. The central concept of this body of work is compensating for field-of-view limitations in various sensing modes, which allows for each mode to intelligently complement the others in an integrated system. However, as mentioned in Section 4.7. above, this core concept is expressed and implemented separately in two areas of the process. The primary mechanism is the parameter adaptivity, embodied by the optimization problem, which uses predicted information statistics to select each mode's next operation regime. The FOV penalty is introduced into the information cost function through an additional measurement covariance term. Covariance control in the state update & association step forms the second pillar of the algorithm, embodied by inclusion of the measurement covariance penalty estimated at the time of the update.

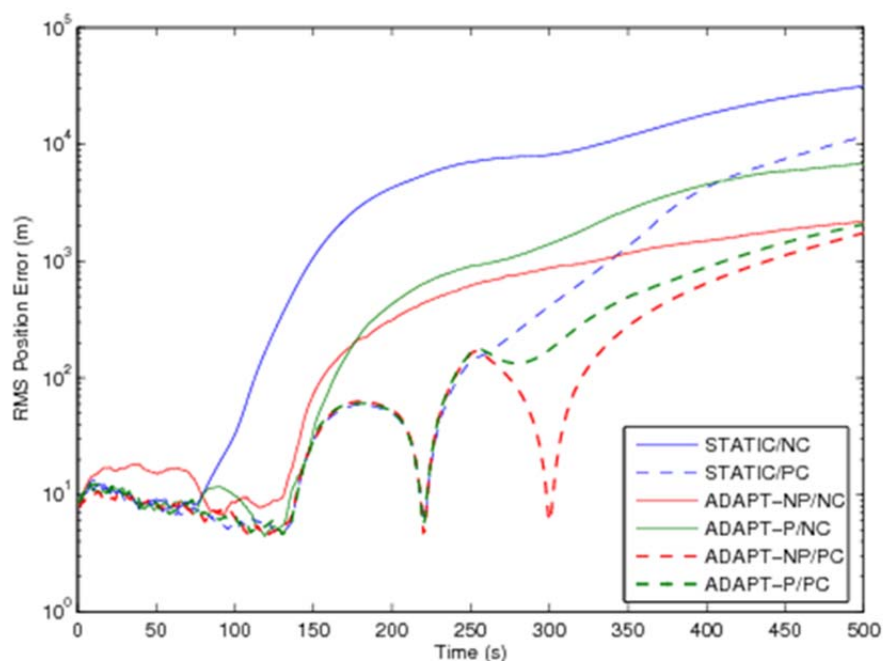
So far, we have usually compared the fully adaptive/compensated method to a non-adaptive/non-compensated method. In order to determine the overall benefit of each aspect, we must consider other combinations of adaptivity & compensation. To this end, four additional algorithms (labels) are considered:

- (1) static mode parameters, penalized covariance compensation (STATIC/PC)
- (2) adaptive parameter selection without FOV cost penalty, no compensation (ADAPT-NP/NC)
- (3) adaptive parameter selection with FOV cost penalty, no compensation (ADAPT-P/NC)
- (4) adaptive parameter selection without FOV cost penalty, penalized compensation (ADAPT NP/PC).

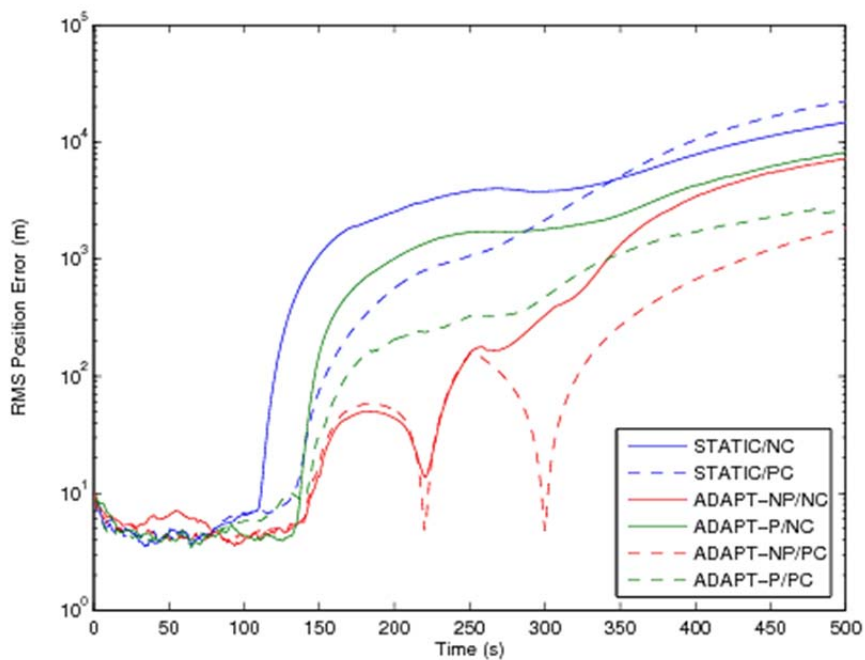
Each of the 6 methods considered, including the previously considered completely non-adaptive (STATIC/NC) and the proposed fully adaptive (ADAPT-P/PC) methods, have been analyzed on both tracking and association grounds, since both are critical to effective integration.

The scenario presented is the so-called “challenge” target scenario in the subsection above: two targets maneuvering away from each other in coordinated turns, tracked by two multimode RF/EO sensors initialized with identical settings over 500 seconds. Coordinated turns and extended motion away from the initial FOVs provide a challenge to the tracking algorithms, requiring appropriate control sequences to use both sensing modes to their maximum potential. Results were obtained from a series of 200 Monte Carlo trials.

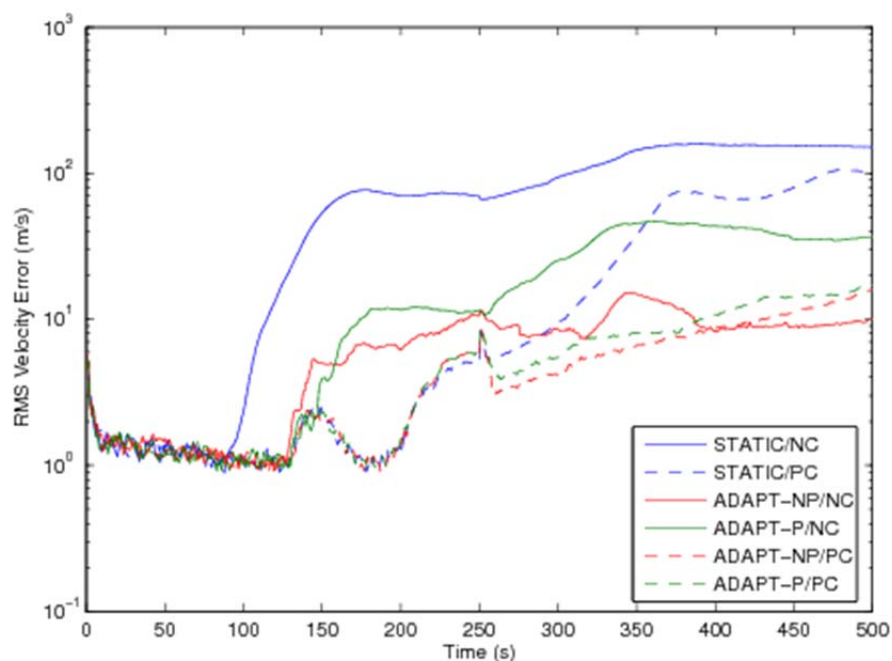
First, the algorithms are compared on tracking error. Figure 26 and Figure 27 show the root-mean-square error of each of the above-described methods for position of Target 1 and 2, respectively. Figure 28 and Figure 29 show the root-mean-square error for velocity of Targets 1 and 2. All of the methods experienced divergence toward the end of the tracking period, which is generally expected for the MSJPDA implementations employed here, as well as in situations employing the approximation objective. As has been shown in previous works, the completely non-adaptive STATIC/NC method performs poorly for both targets, achieving the largest error and diverging the quickest in each case. In most cases, it is clear from the figures that tracking performance is improved by including the field-of-view covariance penalty in the update (the “/PC” subset of algorithms, as shown by the dotted lines), especially for the first target. Adding adaptation only further improves performance, as both ADAPT-NP/PC and the previously proposed ADAPT-P/PC both consistently attain the lowest error by the end of the tracking period and handle the maneuver (starting around 150 seconds) gracefully. The fact that a *non*-penalized cost function provides the best tracking performance may seem surprising at first; however, the compensation in the update step essentially overestimates the optimized variance resulting in lower error. Furthermore, the difference between ADAPT-NP/PC and ADAPT-P/PC is relatively small, and ADAPT-P/PC remains fairly unchanged when ADAPT-NP/PC starts to “go off the rails.”



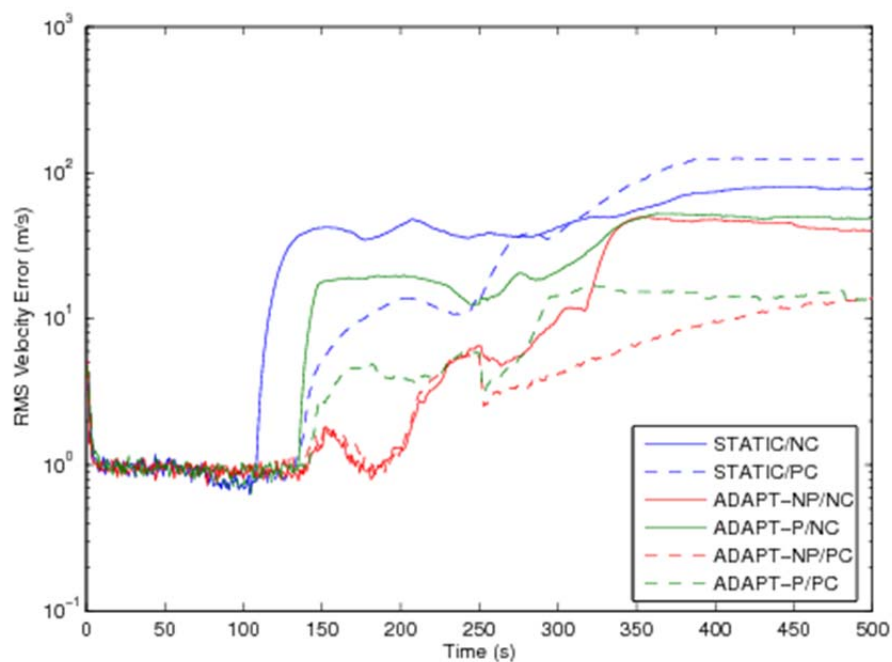
**Figure 26. RMS Position Error Comparison for Target 1**



**Figure 27. RMS Position Error Comparison for Target 2**

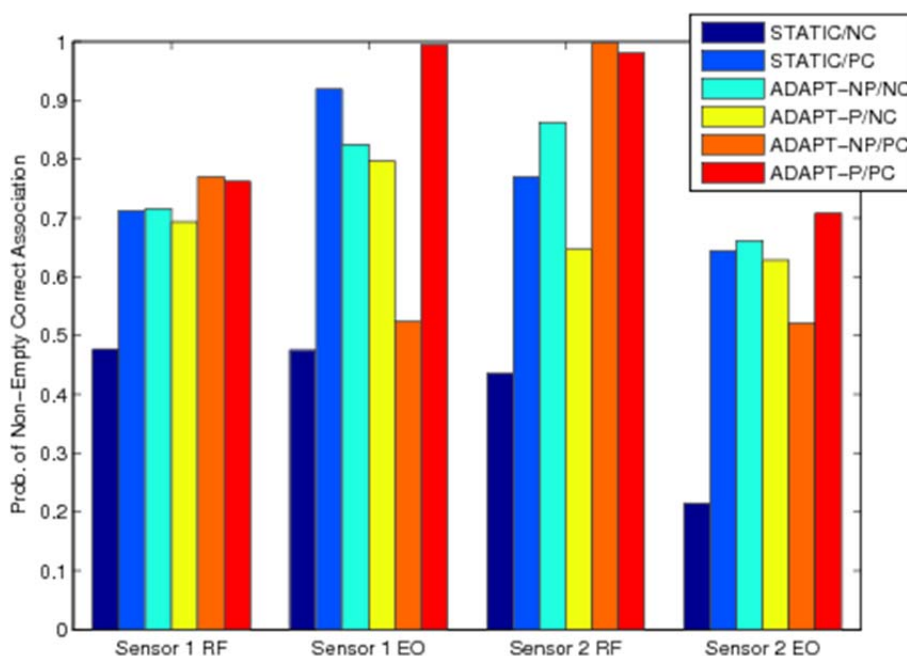


**Figure 28. RMS Velocity Error Comparison for Target 1**



**Figure 29. RMS Velocity Error Comparison for Target 2**

As mentioned above, association performance is also an important component of evaluating these algorithms. The metric considered here is probability of “non-empty” correct association, *i.e.*, correct association when a measurement is made, which is also the metric considered in Section 4.8.2. above. Figure 30 shows this metric for each mode of the two sensors in the scenario. Once again, STATIC/NC is the across-the-board loser, failing to break even 50 percent across all modes. The rest of the algorithms demonstrate very similar performance for the RF mode, with the ADAPT/PC algorithms as clear leaders, exceeding the next best methods by at least ten percent. The EO mode, however, is a different story altogether. STATIC/PC emerges as a strong contender in this respect, primarily because the penalized covariance results in a more precise validation region in spite of the lack of adaptivity. Furthermore, the primary flaw in ADAPT-NP/PC is exposed: the better tracking performance is obtained at the expense correctly associating measurements. The low association performance also comes with many periods of effective inactivity, where no measurement is detected at all. Examinations of sample policies reveal this is the result of frequent switching between high- and low-visibility regimes. In contrast, the proposed ADAPT-P/PC algorithm complements its very good tracking performance with exceptional association performance, outpacing ADAPT-NP/PC handily. While analysis shows that this algorithm also has many non-measurement periods for the EO mode, it is not nearly as much as ADAPT-NP/PC. This results in a nearly 99 percent correct association rate for sensor 1 -- when it makes a measurement, it makes it count. Thus, the proposed ADAPT-P/PC has been shown to provide good balance between and overall effectiveness in both tracking and association.



**Figure 30. Probability of Non-Empty Association Comparison**

## 4.9. Conclusions

We demonstrated that the field-of-view constraint is a necessary consideration in any multi-target tracking system, but especially for those that use multi-modal sensing mechanisms. In order to effectively reap the benefits of the multiple modalities, a comprehensive understanding of their limitations is essential. Additionally, we devised an appropriate modeling mechanism that accounted for the FOV effect in a simple, but robust fashion. The variance penalty captures the essential feature of a missed measurement due to a target being out-of-view – increased estimation uncertainty – while also being simple to include in the estimation procedure and the parameter optimization process.

The compensation and optimization scheme is also a significant advance. We showed that by keeping the multi-modal sensing concept in mind resulted in a straightforward means to mitigate the FOV effect in the state estimation procedure and derive effective objective functions and constraints for parameter selection. This algorithm was shown to be beneficial in terms of tracking performance in both ideal and cluttered conditions, for random and maneuvering targets. Furthermore, we demonstrated that the algorithm resulted in improved data association over non- and lesser-adaptive methods with little modification from the perfect detection scheme. Finally, we demonstrated that both aspects of this method, penalized feedback compensation and penalized parameter adaptation, are necessary to completely mitigate the FOV effect. While the adaptation without penalization can provide marginally better tracking performance, the fully penalized method made up for the minimal tracking deficit by improving association capabilities significantly. This demonstrates an inherent tradeoff between tracking and association performance, one that our method deftly balances.

The key advantage of multi-modal sensors is the associated increase in degrees of freedom available for target detection, tracking and association. The ability to exploit this additional flexibility requires the use of adaptive or "cognitive" sensing techniques that take the current operational scenario into account and adjust the sensor parameters accordingly. In the RF-EO sensors studied as part of this effort, these parameters include the EO field-of-view (FOV) and resolution, as well as the radar bandwidth and pulse repetition frequency. If these parameters are not carefully chosen, new targets will be missed, tracking errors will diverge and track associations will be lost. Our contribution is an automated method for adapting the RF-EO multi-model sensor operation so that it minimizes the likelihood of such events. Rather than just using a brute-force approach where one steps through all possible combinations of the sensor parameters (clearly not a realistic solution when several multi-model sensors are present with multiple degrees of freedom), the sensor's operation is adjusted according to the currently detected targets and their associated tracks using an automated procedure that takes into account the competing objectives associated with tracking multiple maneuvering targets. The system operation then alternates between this target-scenario-optimized behavior and a wide-FOV search for new targets. Our simulation results indicate that order-of-magnitude improvements and better are possible using our adaptive approach compared with static, non-adaptive sensor settings. These results are consistent with other work on cognitive radar, and provide strong motivation for their continued study.

## 5. Tracking and Waveform Agility

### 5.1. Methods, Assumptions, and Procedures

Sensing systems with multiple modalities and multiple interactive processing capabilities are expected to achieve high target tracking performance. Specifically, reduced detection false alarm rate, and thus reduced target tracking error, result from combining kinematic feature information from radio frequency (RF) radar sensors with physical feature information from electro-optical (EO) sensors in an asymmetric multi-modal tracking system [73]. For example, an RF radar in ground moving target indicator (GMTI) mode or high-range resolution radar (HRRR) mode uses a wide field of view (FOV) to track moving targets whereas a modally asymmetric EO sensor in staring imaging mode uses a narrow FOV to identify stationary targets [74]. Adopting the joint asymmetric RF-EO modality can be further exploited in complex challenging sensing scenarios, including low signal-to-noise ratio (SNR), high clutter or interference, weather environmental conditions, and light intensities. Multi-modality sensing can be further integrated with intelligent processing techniques, such as radar waveform parameter and EO resolution parameter configuration or advanced Bayesian techniques to increase tracking performance.

Low observable targets require many unthresholded measurements to increase their probability of detection. This is because, as the strength of the received signal depends on the target's radar cross section (RCS) and the distance from the radar to the target, and radars with limited aperture size, transmit power, or duty cycle have low probability of detecting smaller objects from great distances, thresholding low SNR targets forces a trade-off between low probability of detection (PD) and high probability of false alarm (PFA). One technique that uses unthresholded matched filter measurements to estimate the target state is the track-before-detect (TBD) algorithm, implemented using particle filtering (PF-TBD) [76, 77]. This is an approach that performs target tracking while detecting, as the tracking iterations are initiated by a set of possible tracks that depend on all available matched filter measurements. In this work, RF-EO multi-modal low SNR tracking of targets with different energy returns was investigated using the PF-TBD approach. In a classical target tracking processing problem, detection is first performed by matched filtering or thresholding the RF-EO measurements collected at each time step. However, as returns from different scenes may experience different energy levels, this can lead to a loss of information in the measurements as it would be difficult to find an appropriate threshold level to use in all possible scenes. This can be avoided using the PF-TBD as it only uses unthresholded measurements.

The PF-TBD algorithm is expected to improve sensing performance in low SNR scenarios. However, as a tradeoff, the volume of data gathered by sensors often places an overwhelming demand on signal processing and results in redundant resource consumptions, especially when incorporated with EEF. On the other hand, adapting the transmit waveform can contribute to more accurate state estimates. Therefore, the ability to intelligently direct sensor configurations have a significant impact on the performance of a sensing system [78]. Dynamic waveform adaptation provides a sensing methodology to design the next transmitted waveform to optimize the tracker's requirements. The system level optimization yields better tracking performance than optimizing the sensors and tracker independently [79, 80]. In this work, an agile-waveform design scheme has been integrated into the PF-TBD algorithm, which can configure the

transmitted waveform on a pulse-to-pulse basis to obtain the information that optimally improves the tracker's performance according to requirements. The performance criteria in the is the predicted mean-squared error, which is approximately evaluated using the Cramér-Rao lower bound (CRLB) at each time step [81, 82, 84].

Although RF and EO sensors observe the same scene at the same time, the measurements are assumed independent from RF cell to cell and EO frame to frame for ease of processing. However, unless the measurements are asynchronous, this assumption is not very realistic. When the measurement noise in the range-Doppler plane and the azimuth plane are assumed dependent and with unknown correlation parameters, the challenge is how to appropriately process them. In this work, the exponentially embedded family (EEF) estimation method is adopted and integrated into the PF-TBD algorithm to approximate the unknown probability density function (PDF) of the dependent measurements [85]. The method asymptotically approximates an unknown PDF with the exponential family PDF that is closest to it using the Kullback-Leibler (KL) divergence measure [80, 82].

When thresholding is used, increasing the threshold to increase the probability of detection requires more measurements; but that can cause an increase in the probability of false alarm. However, the additional measurements can exponentially increase the computational complexity of multiple target tracking (MTT) algorithms, such as the joint probabilistic data association (JPDA) and the multiple hypothesis tracking [87]. Tracking a time-varying number of targets under low SNR and high clutter is still considered a difficult problem in radar. The PF-TBD has been extended to tracking two targets, when the second target spawns from the first target [76]; a multi-camera multi-target tracking algorithm based on particle clustering [88]. The recursive TBD algorithm is generalized to track multiple targets by estimating the joint posterior density of their state vector while keeping track of targets entering and leaving the FOV at any time step. The new multiple-mode multiple-target PF-TBD algorithm is based on estimating the posterior probability density of the target states under different modes in order to consider all possible target existence combinations at each time step [89].

The same trend towards increased processing capabilities has led to the development of multiple target tracking algorithms for a known number of targets, such as the joint probabilistic data association (JPDA) [90] and multiple hypothesis tracker (MHT). Recently, a new tracking algorithm, the probability hypothesis density filter (PHDF), is being used for multiple target tracking [91]. The PHDF is based on the theory of random finite sets to simultaneously track multiple targets, propagating a first-order statistical moment of the random finite set. Unlike the JPDA and MHT, the PHDF does not need to know a priori the number of targets, as the number of targets is estimated as part of the tracking algorithm. In this work, the multiple target tracking capabilities of the PHDF is combined with dynamic waveform selection based on the predicted mean-squared error (MSE). A time varying number of targets are tracked, and the predicted MSE of individual targets is used to select the waveform used at the next time step, resulting in a single integrated multiple-target, waveform-agile tracker [92].



## 5.2. Multi-Modal Track-Before-Detect for Low Observable Targets

We developed the particle filter (PF) based recursive track-before-detect (TBD) approach for joint asymmetric multi-modal radio frequency (RF) radar and electro-optical (EO) sensing to avoid loss of information caused by matched filter thresholding at low SNR [81]. This work was performed in collaboration with the PI, Muralidhar Rangaswamy.

### 5.2.1. Target State Model for Track-Before-Detect

We consider an RF-EO asymmetric multi-modal sensing system for tracking a point target moving in a two-dimensional (2-D) space under low SNR conditions. The unknown target state vector  $\mathbf{x}_k = [x_k \dot{x}_k y_k \dot{y}_k]^T$ , where  $T$  denotes vector transpose, at time step  $k$  consists of the 2-D Cartesian coordinates of the target's position  $(x_k, y_k)$  and velocity  $(\dot{x}_k, \dot{y}_k)$ . The state evolution, assuming constant velocity model, is

$$\mathbf{x}_k = \mathbf{F} \mathbf{x}_{k-1} + \mathbf{w}_k \quad (191)$$

where

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (192)$$

$\Delta t$  is the time between steps, and  $\mathbf{w}_k$  is the modeling error random process, which is assumed zero-mean Gaussian with noise intensity  $q$  and covariance matrix

$$\mathbf{Q} = q \begin{bmatrix} (\Delta t)^3/3 & 0 & (\Delta t)^2/2 & 0 \\ 0 & (\Delta t)^3/3 & 0 & (\Delta t)^2/2 \\ (\Delta t)^2/2 & 0 & \Delta t & 0 \\ 0 & (\Delta t)^2/2 & 0 & \Delta t \end{bmatrix}. \quad (193)$$

Depending on the presence of a target in a scene, the TBD algorithm describes these state transitions as a jump Markov system with Markov matrix  $\Xi_k = \begin{bmatrix} 1 - \Pr_k^{(b)} & \Pr_k^{(b)} \\ \Pr_k^{(d)} & 1 - \Pr_k^{(b)} \end{bmatrix}$ . The probability of a target entering and leaving the scene at time  $k$  are, respectively, given by  $\Pr_k^{(b)} = \Pr(E_k = 1 | E_{k-1} = 0)$  and  $\Pr_k^{(d)} = \Pr(E_k = 0 | E_{k-1} = 1)$ . Here, the binary variable  $E_k$  is used to formulate the presence or absence of a target at time step  $k$ . Specifically, it is assumed that the target is present if  $E_k = 1$  and absent if  $E_k = 0$ .

### 5.2.2. RF-EO Multi-Modal Measurement Model for Track-Before-Detect

The symmetric or asymmetric fusion of radio frequency (RF) radar and electro-optical (EO) sensor measurements for joint target tracking is an example of a multi-modal sensing platform. Radars can be used to detect and track targets moving with radial velocity as they provide range and range-rate measurements; EO sensors can be used to provide azimuth and elevation angle information to detect stationary targets. Conventionally, in order to use the RF-EO sensor measurements for tracking, matched filtering and thresholding to decide the presence of a target is necessary. Using the TBD algorithm, all unthresholded RF-EO sensor measurements need to be processed; all range-Doppler plane outputs are required to form the RF measurements and all EO sensor correlation plane outputs are required to form the EO measurements. As a result, the TBD algorithm keeps all potentially useful information and fuses data across the entire range-Doppler plane (for the RF) and angle measurement plane (for the EO). It must be noted that this process can be computationally intensive if the dimensionality of the measurements is high.

#### 5.2.2.1. Radar Sensor Measurement Model

For the RF sensor, each range-Doppler resolution cell is assumed to provide a matched filter output, and the 2-D range-Doppler plane is divided into  $L = A \times B$  rectangular resolution cells. The resolution cell center is denoted by  $(r_l, \dot{r}_l)$ ,  $l = 1, \dots, L$ . The RF sensor measurement at the  $l$ th cell center at time step  $k$ , assuming that  $E_k = 1$ , is given by,

$$z_{k,l}^{(\text{RF})} = h_{k,l}^{(\text{RF})}(\mathbf{x}_k) + v_{k,l} = I^{(\text{RF})} \text{AF}_s \left( \frac{1}{2c}(r_l - r_k), \frac{2f_c}{c}(\dot{r}_l - \dot{r}_k) \right) + v_{k,l} \quad (194)$$

where  $I^{(\text{RF})}$  is the radar return energy intensity,  $\text{AF}_s(\tau, \nu)$  is the ambiguity function (AF) of the transmitted signal  $s(t)$ ,  $c$  is the velocity of the electromagnetic waves in air,  $v_{k,l}$  is additive noise at the  $l$ th resolution cell, and  $f_c$  is the carrier frequency. The relationship between range  $r_k$  and range-rate  $\dot{r}_k$  and the target position and velocity in (191) at time step  $k$  can be obtained as

$$\begin{aligned} r_k &= \sqrt{(x_k - x_{\text{sr}})^2 + (y_k - y_{\text{sr}})^2} \\ \dot{r}_k &= \frac{1}{r_k} ((x_k - x_{\text{sr}})\dot{x}_k + (y_k - y_{\text{sr}})\dot{y}_k) \end{aligned} \quad (195)$$

where  $(x_{\text{sr}}, y_{\text{sr}})$  is the 2-D (fixed) location coordinates of the sensor.

The likelihood PDF of the measurement vector  $\mathbf{z}_k^{(\text{RF})} = [z_{k,1}^{(\text{RF})} z_{k,2}^{(\text{RF})} \dots z_{k,L}^{(\text{RF})}]^T$ , is given by  $p(\mathbf{z}_k^{(\text{RF})} | \mathbf{x}_k, E_k = 1) = p(\mathbf{z}_k^{(\text{RF})} | \mathbf{x}_k)$  or  $p(\mathbf{z}_k^{(\text{RF})} | \mathbf{x}_k, E_k = 0) = p(\mathbf{z}_k^{(\text{RF})})$ , depending on whether the target is present or not. When the target is not present,  $E_k = 0$  and  $z_{k,l}^{(\text{RF})} = v_{k,l}$ . The measurement likelihood ratio for detection based on all unthresholded measurements in all cells is given by:

$$\Lambda_{\text{RF}}(z_{k,l}^{(\text{RF})} | \mathbf{x}_k) = p_{s+N}(z_{k,l}^{(\text{RF})} | \mathbf{x}_k) / p_N(h_{k,l}^{(\text{RF})}(\mathbf{x}_k)) = \exp(-h_{k,l}^{(\text{RF})}(\mathbf{x}_k)(h_{k,l}^{(\text{RF})}(\mathbf{x}_k) - 2z_{k,l}^{(\text{RF})}) / \sigma_{\text{RF}}^2) \quad (196)$$

If we assume that the measurement noise at each cell is independent, the likelihood ratio of the whole RF measurement vector is

$$\Lambda_{\text{RF}}(\mathbf{z}_k^{(\text{RF})} | \mathbf{x}_k, E_k) = \begin{cases} \prod_l \Lambda_{\text{RF}}(z_{k,l}^{(\text{RF})} | \mathbf{x}_k), & E_k = 1 \\ 1, & E_k = 0 \end{cases} \quad (197)$$

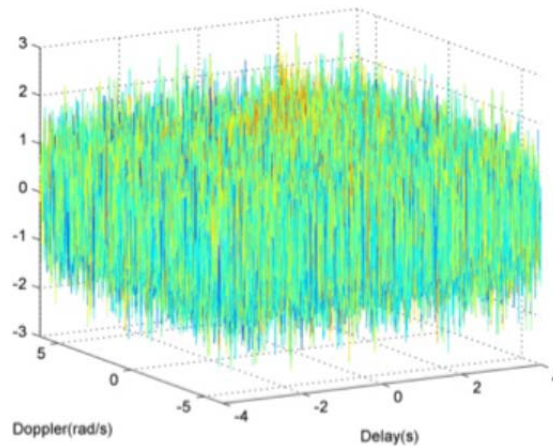
In order to demonstrate the need to keep all unthresholded measurements in the AF plane, a linear frequency-modulated (LFM) chirp transmit signal is used, given by

$$s(t) = 1/(\pi T^2)^{1/4} \exp(-t^2/(2T^2)) \exp(jbt^2) \quad (198)$$

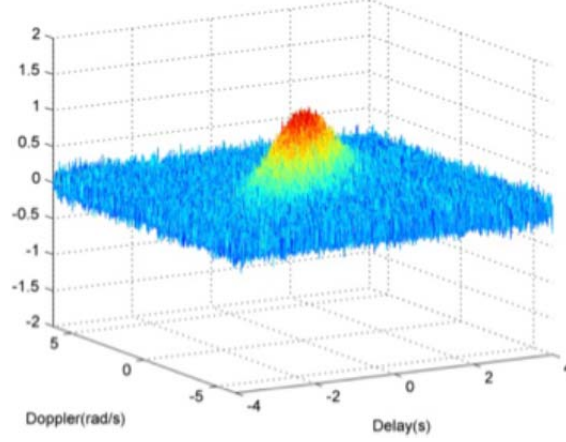
where  $b$  is the frequency-modulation (FM) rate and  $T$  is the standard deviation of the Gaussian envelope. The magnitude of its AF is given by

$$\square \quad | \text{AF}_s(\tau, \nu) | = \left| \int_{-\infty}^{\infty} s(t) s^*(t + \tau) e^{-j2\pi\nu t} dt \right| = \exp\left[-\left(\tau^2/(4T^2) + \pi^2 T^2 (\nu + b\tau/\pi)^2\right)\right] \quad (199)$$

Note that if delay  $\tau_k$  and Doppler  $\nu_k$  estimates are obtained at time step  $k$  from the AF, then the target range  $r_k$  and range-rate  $\dot{r}_k$  can be obtained as  $r_k = c \tau_k / 2$  and  $\dot{r}_k = c \nu_k / (2f_c)$ , respectively. Figure 31 and Figure 32 show the LFM AF at 40 dB and 6 dB SNR, respectively. Figure 32 clearly shows that, at low SNR, thresholding would definitely result in loss of information since a single delay-Doppler point cannot be identified that can be associated with the target.



**Figure 31. Ambiguity Function Measurement at 6 dB SNR.**



**Figure 32. Ambiguity Function Measurement at 40 dB SNR.**

#### 5.2.2.2. EO Sensor Measurement Model

For the EO sensor, it is similarly assumed that each resolution cell provides a matched filter output. Considering only azimuth angle resolution cells, the 1-D (along azimuth) angular AF is divided into  $M$  resolution cells, each of length  $\theta_m$ , where  $m = 1, \dots, M$ . At a given time  $k$ , if  $E_k = 1$ , the EO angle measurement obtained at the  $m$ th resolution cell is given by,

$$z_{k,m}^{(\text{EO})} = h_{k,m}^{(\text{EO})} + u_{k,m} = I^{(\text{EO})} A_d(\theta_m - \theta_k, 0) + u_{k,m} \quad (200)$$

where  $I^{(\text{EO})}$  is the EO return energy intensity,  $A_d(\theta, \phi)$  is the angular AF of the transmitted signal  $d(t)$ ,  $\theta_k = \arctan(y_k/x_k)$  is the angle of the target, and  $u_{k,m}$  is additive noise at the  $m$ th resolution cell. Depending on the presence of the target, the measurement likelihood PDF is

- $p(\mathbf{z}_k^{(\text{EO})} | \mathbf{x}_k, E_k = 1) = p(\mathbf{z}_k^{(\text{EO})} | \mathbf{x}_k)$  or  $p(\mathbf{z}_k^{(\text{EO})} | \mathbf{x}_k, E_k = 0) = p(\mathbf{z}_k^{(\text{EO})})$ , where the measurement
- vector is  $\mathbf{z}_k^{(\text{EO})} = [z_{k,1}^{(\text{EO})} \ z_{k,2}^{(\text{EO})} \ \dots \ z_{k,M}^{(\text{EO})}]^T$ . When  $E_k = 0$  (target not present), that is,  $z_{k,m}^{(\text{EO})} = u_{k,m}$ . Then the likelihood ratio of the whole RF measurement vector is

$$\Lambda_{\text{EO}}(\mathbf{z}_k^{(\text{EO})} | \mathbf{x}_k, E_k) = \begin{cases} \prod_m \exp(-h_{k,m}^{(\text{EO})}(\mathbf{x}_k)(h_{k,m}^{(\text{EO})}(\mathbf{x}_k) - 2z_{k,m}^{(\text{EO})})/\sigma_{\text{EO}}^2) & E_k = 1 \\ 1 & E_k = 0 \end{cases} \quad (201)$$

#### 5.2.2.3. RF-EO TBD Implementation Using Particle Filtering

The TBD algorithm can be implemented using a PF. The PF will use a measurement equation that is based on all unthresholded measurements from the range-Doppler plane and the angle plane. If the measurements from the RF-EO sensors are assumed independent, then the weight  $w_k^{(n)} = \Lambda_{\text{RF}}(\mathbf{z}_k^{(\text{RF})} | \mathbf{x}_k, E_k^{(n)}) \Lambda_{\text{EO}}(\mathbf{z}_k^{(\text{EO})} | \mathbf{x}_k, E_k^{(n)})$  for the  $n$ th particle depends on the measurement likelihood functions in Equations (197) and (201), respectively.

### 5.3. Waveform-Agile Multi-Modal TBD

We developed a waveform optimization approach to allow for adaptive transmit waveform selection under low SNR conditions. The approach is integrated into the TBD to select the transmit waveform parameters at each time step by minimizing the predicted mean-squared estimation error of the state of the target [81, 82, 83]. This work was performed in collaboration with the PI, Muralidhar Rangaswamy. In our previous work, waveform agility was achieved under the assumption that the measurement noise covariance matrix can be approximated by the CRLB covariance matrix. At high SNR values, this assumption is valid because a single range and range-rate measurement, obtained by thresholding the AF of the received signal, can provide sufficient information to estimate the target state. In particular, the CRLB characterization can be obtained directly from the curvature of the AF peak at the origin of the delay-Doppler plane [78]. The high SNR assumption used in the waveform design technique does not hold under low SNR environments. In such cases, thresholding low SNR targets forces a trade-off between a low PD and a high PFA. Thus, low observable targets require *unthresholded* measurements that are below the target detection threshold, to increase their probability of detection. The TBD approach uses the unthresholded measurements and it tracks while detecting, as the tracking iterations are initiated by a set of possible tracks that depend on all available AF measurements. The TBD is computationally feasible as it can be implemented using a particle filter (PF) [75].

We consider a library of  $L$  waveforms,  $s(t; \beta_l)$  with parameter set  $\beta_l = [T_l b_l]$ ,  $l = 1, \dots, L$ . If the waveforms are LFM, for example, then  $b_l$  is the FM rate and  $T_l$  is the standard deviation of the Gaussian envelope of the  $l$ th LFM chirp. Using the PF-TBD, we compute the predicted covariance matrix, assuming the  $l$ th LFM chirp was transmitted using

$$\tilde{\mathbf{P}}_{k+1, \beta_l} = \left( \tilde{\mathbf{J}}_{k+1, \beta_l} + [\mathbf{F} \mathbf{M}_k \mathbf{F}^T + \mathbf{Q}]^{-1} \right)^{-1} \quad (202)$$

where  $\mathbf{M}_k = \sum_n w_k^{(n)} (\mathbf{x}_k^{(n)} - \hat{\mathbf{x}}_k)(\mathbf{x}_k^{(n)} - \hat{\mathbf{x}}_k)^T$ , and  $\hat{\mathbf{x}}_k$  is the state estimate obtained using particles  $\mathbf{x}_k^{(n)}$  at time step  $k$  and  $w_k^{(n)}$  are the corresponding weights. The expected information matrix can be approximated as

$$\tilde{\mathbf{J}}_{k+1, \beta_l} \approx \sum_n \tilde{\mathbf{H}}_{k+1, \beta_l}^{(n)T} \mathbf{R}_k^{-1} \tilde{\mathbf{H}}_{k+1, \beta_l}^{(n)} \quad (203)$$

where  $\tilde{\mathbf{H}}_{k+1, \beta_l}^{(n)} = \Delta \tilde{\mathbf{x}}_{k+1}^{(n)} h_{k,l}(\tilde{\mathbf{x}}_{k+1}^{(n)})$  and  $h_{k,l}(\tilde{\mathbf{x}}_{k+1}^{(n)})$  depends on the RF or EO sensor as defined in Equations (194) and (198), respectively. The operator  $\Delta$  computes the gradient with respect to each of the target state parameters and  $\tilde{\mathbf{x}}_{k+1}^{(n)} = \mathbf{F} \mathbf{x}_k^{(n)}$  is the forward predicted  $n$ th particle.

The waveform agile algorithm selects the waveform with parameter set  $\hat{\beta}$  that minimizes the weighted predicted MSE at time step  $(k+1)$ ,

$$s_{k+1}(t; \hat{\beta}) = \arg \left\{ \min_{\beta_l} \left[ \mathbf{g}^T \text{diag}(\tilde{\mathbf{P}}_{k+1, \beta_l}) \right] \right\}. \quad (204)$$

The weight vector  $\mathbf{g} = [g_x g_{\dot{x}} g_y g_{\dot{y}}]^T$  is used to scale the relative importance of the position and velocity estimation errors.

#### 5.4. Statistical Dependent RF-EO Measurements

We proposed the use of the embedded exponential family (EEF) approach [85] to approximate the joint PDF of the dependent RF-EO measurements in order to further improve the PF-TBD tracker performance [81, 82]. This work was performed in collaboration with Co-PIs Steven Kay and Muralidhar Rangaswamy.

In the previous sections, we have assumed that measurement noise realizations from each range-Doppler cell and angle resolution cell are independent. However, unless the measurements are asynchronous, they should be considered statistical dependent and the unknown measurement noise joint PDF needs to be estimated for use in the PF-TBD algorithm. We form the general EEF dependence term vectors  $\boldsymbol{\eta} = [\eta_1 \eta_2 \cdots \eta_L]^T$  and  $\boldsymbol{\varsigma} = [\varsigma_1 \varsigma_2 \cdots \varsigma_M]^T$  by concatenating the range-Doppler and angle resolution cells, respectively. Using the EEF approach [85, 86], we form the particle filter weights as  $w_k^{(n)} = 1$  when  $E_k^{(n)} = 0$ , and when  $E_k^{(n)} = 1$ , we recomputed them as  $w_k^{(n)} = \Lambda_{\text{RF}}(\mathbf{z}_k^{(\text{RF})} | \mathbf{x}_k, E_k^{(n)}, \boldsymbol{\eta}) \Lambda_{\text{EO}}(\mathbf{z}_k^{(\text{EO})} | \mathbf{x}_k, E_k^{(n)}, \boldsymbol{\varsigma})$ . This simplifies to

$$w_k^{(n)} = \exp \left[ \sum_l \eta_l \ln(p_{s+N}(\mathbf{z}_{k,l}^{(\text{RF})} | \mathbf{x}_k^{(n)}) / p_N(\mathbf{z}_{k,l}^{(\text{RF})})) + \sum_m \varsigma_m \ln(p_{s+N}(\mathbf{z}_{k,m}^{(\text{EO})} | \mathbf{x}_k^{(n)}) / p_N(\mathbf{z}_{k,m}^{(\text{EO})})) - G(\boldsymbol{\eta}, \boldsymbol{\varsigma}) + \ln(p_N(\mathbf{z}_k^{(\text{RF})}, \mathbf{z}_k^{(\text{EO})})) \right] \quad (205)$$

where

$$G(\boldsymbol{\eta}, \boldsymbol{\varsigma}) = \sum_l (\eta_l^2 - \eta_l) (\bar{z}_{k,l}^{(\text{RF})})^2 / (2\sigma_{\text{RF}}^2) + \sum_m (\varsigma_m^2 - \varsigma_m) (\bar{z}_{k,m}^{(\text{EO})})^2 / (2\sigma_{\text{EO}}^2) \quad (206)$$

and

$$p_N(\mathbf{z}_k^{(\text{RF})}, \mathbf{z}_k^{(\text{EO})}) = \Pi_l (1/\sqrt{2\pi\sigma_{\text{RF}}^2}) \exp(-(z_{k,l}^{(\text{RF})})^2 / (2\sigma_{\text{RF}}^2)) \Pi_m (1/\sqrt{2\pi\sigma_{\text{EO}}^2}) \exp(-(z_{k,m}^{(\text{EO})})^2 / (2\sigma_{\text{EO}}^2)) \quad (207)$$

The maximum likelihood estimator of  $\boldsymbol{\eta}$  and  $\boldsymbol{\varsigma}$  is found by maximizing

$$\frac{1}{\sigma_{\text{RF}}^2} \sum_l \eta_l z_{k,l}^{(\text{RF})} \bar{z}_{k,l}^{(\text{RF})} - \frac{1}{\sigma_{\text{RF}}^2} \sum_l \eta_l^2 (\bar{z}_{k,l}^{(\text{RF})})^2 \quad (208)$$

and

$$\frac{1}{\sigma_{\text{EO}}^2} \sum_m \varsigma_m z_{k,m}^{(\text{EO})} \bar{z}_{k,m}^{(\text{EO})} - \frac{1}{\sigma_{\text{EO}}^2} \sum_m \varsigma_m^2 (\bar{z}_{k,m}^{(\text{EO})})^2 \quad (209)$$

under the constraint  $0 < \eta_l, \varsigma_m < 1$ , which guarantees that  $G(\boldsymbol{\eta}, \boldsymbol{\varsigma})$  is finite. Note that  $\bar{z}_{k,l}^{(\text{RF})}$  and  $\bar{z}_{k,m}^{(\text{EO})}$  are the noiseless contributions of the RF-EO measurements given the particle  $\mathbf{x}_k^{(n)}$ .

## 5.5. Multiple Target Tracking

We investigated approaches for multiple target tracking for different scenerios, including multi-modal sensors with multiple low observable targets and adaptive multiple target tracking when the number of targets at each step is unknown. This work was performed in collaboration with the PI, Muralidhar Rangaswamy.

### 5.5.1. Rao-Blackwellized Particle Filter for an Unknown Number of Targets

#### 5.5.1.1. Joint RF-EO Multi-Target Tracking

We investigated the joint multi-modal operation of the asymmetric character of the fields-of-view of radar and EO sensors for multi-target tracking applications. We consider pulsed-Doppler radar and EO sensors as an attractive sensor combination as the radar provides range and range-rate measurements to detect targets with radial velocity and the EO sensor provides azimuth and elevation angles information to detect stationary targets.

We formulate the tracking problem using a target state equation model defined to describe point targets moving in 3-D, and RF-EO measurement models specified for target-originated and clutter measurements. The complete formulation includes probabilistic stochastic process models for target states, data associations, and processes for targets entering or leaving a scene. We solve the tracker of an unknown number of targets using the Rao-Blackwellized particle filter (RBPF) algorithm [93, 94] based on sequential Monte Carlo (SMC) sampling techniques. The RBPF decomposes a filtering problem that would require Monte Carlo sampling into two filtering problems: (a) one that can be solved in closed form; and (b) a lower dimensionality one than the original that would require Monte Carlo sampling but would not be as computationally intensive as the original one. Also, solving some of the equations in closed form instead of using Monte Carlo sampling for all the equations can be shown to produce estimators with lower variance. Considering a system with state vector  $S_k$  and measurement vector  $\zeta_k$  at time step  $k$ , the state space model can be given by the density functions  $p(S_k | S_{k-1})$  and  $p(\zeta_k | S_k)$ . If we partition the state vector into  $S_k = [\chi_k^T \ \lambda_k^T]^T$ , the state space model can be re-written as  $p(\chi_k | \chi_{k-1}, \lambda_{k-1}, \lambda_k)$ ,  $p(\lambda_k | \chi_{k-1}, \lambda_{k-1})$ , and  $p(\zeta_k | \chi_k, \lambda_k)$ . If  $p(\chi_k | \chi_{k-1}, \lambda_{k-1}, \lambda_k)$  and  $p(\zeta_k | \chi_k, \lambda_k)$  can be calculated in closed form, we can apply Monte Carlo sampling only to  $\lambda_k$ , with  $\lambda_k$  independent of  $\chi_{k-1}$  so that  $p(\lambda_k | \chi_{k-1}, \lambda_{k-1}) = p(\lambda_k | \lambda_{k-1})$ . In this way, a decrease in the state dimensionality and use of the RBPF results in a largely reduced number of particles, improving the effectiveness of the Monte Carlo sampling and thus leading to an efficient implementation. For the joint RF-EO multi-target tracking, the partitions of  $S_k$  are given by the target position and velocity  $\chi_k$  and by the various association events  $\lambda_k$ . Here, we assume that the transition model for the association events  $p(\lambda_k | \lambda_{k-1})$  is independent of  $\chi_{k-1}$ .

#### 5.5.1.2. Waveform Agile Multi-Modal Design

We proposed a joint multi-modal sensing mode based on using dynamic agility selection to optimize the tracking performance of multiple maneuvering targets. The proposed method jointly

designs waveforms for radar sensing and resolution switching modes for EO sensing, when both sensor measurements experience high false alarm rates. By independently and adaptively configuring the radar waveform parameters and EO resolution parameters, we show that it is possible to increase the joint sensing performance and fully exploit multi-modality.

The approach is based on deriving the posterior CRLB (PCRLB) for single and multiple target tracking using the joint RF-EO sensor system. If  $\hat{x}_k$  is an unbiased estimate of the target state  $x_k$  obtained using RF-EO measurements, then the covariance matrix of  $\hat{x}_k$  has the lower bound  $P_k = E[(\hat{x}_k - x_k)(\hat{x}_k - x_k)^T] \geq J_k^{-1}$ . The Fisher information matrix  $J_k$  can be computed recursively as  $J_{k+1} = Q_k + J_k^{RF/EO} - F^T Q_k^{-1} (J_k + F^T Q_k^{-1} F)^{-1} (Q_k^{-1})^T F$ , where  $F$  and  $Q_k$  are the target state transition and process noise covariance matrices, respectively, and  $J_k^{RF/EO}$  depends on the probabilities of detection and false alarm, validation gate volume, and measurement covariances for the RF and EO sensors. In general, higher PD or lower PFA yield higher information and consequently better tracking performance.

A direct optimization of the error covariance matrix can provide guidelines for waveform design for the RF-EO sensing system to improve tracking performance. However, due to the computational complexity involved, inferences on a waveform design scheme were instead derived from the PCRLB. Specifically, at time step  $k$ , we obtain the target state estimate  $\hat{x}_k$  using the RBPF and then use this estimate to predict the possible states at the next time step and the corresponding distributions using target state models. The predicted target states are then used with the PCRLB to select the sensor parameters. We demonstrated the performance of the proposed adaptive tracking system using numerical simulations.

## 5.5.2. Tracking Low Observable Targets Using PF-TBD

Tracking a time-varying number of targets under the severe conditions of low SNR and high clutter is still considered a difficult problem in radar. We generalized the recursive TBD algorithm to track multiple low observable targets by estimating the joint posterior PDF of their state vector while keeping track of targets entering and leaving the FOV at each time step. The different modes in the multiple target case correspond to the different number of target combinations that may be present in the scene at time steps  $k-1$  and  $k$ . The state vectors corresponding to all the targets in the different modes are then integrated to derive the overall target estimate. Thus, the proposed multiple-mode multiple-target PF-TBD algorithm is based on estimating the posterior probability density of the target states under different modes in order to consider all possible target presence combinations at each time step.

### 5.5.2.1. Multiple Target Model for PF-TBD

Assuming a maximum possible number of  $L$  targets, the combined target state vector is given by  $\mathbf{x}_k = [\mathbf{x}_{k,1}^T \quad \mathbf{x}_{k,2}^T \quad \dots \quad \mathbf{x}_{k,L}^T]^T$ , with the  $l$ th target state vector  $\mathbf{x}_{k,l} = [x_{k,l} \quad \dot{x}_{k,l} \quad y_{k,l} \quad \dot{y}_{k,l} \quad I_{k,l}]^T$ ,  $l = 1, \dots, L$ . The 2-D Cartesian coordinates of the position and velocity of the  $l$ th moving target at time step  $k$  are given by  $(x_{k,l}, y_{k,l})$  and  $(\dot{x}_{k,l}, \dot{y}_{k,l})$ , respectively, and  $I_{k,l}$  is the measurement



intensity level of the  $l$ th target in the associated position. The multiple target state model for all  $L$  targets can be written as

$$\mathbf{x}_k = \begin{bmatrix} \mathbf{F} & \cdots & 0 \\ \vdots & \mathbf{F} & \vdots \\ 0 & \cdots & \mathbf{F} \end{bmatrix} \mathbf{x}_{k-1} + \mathbf{v}_k \quad (210)$$

where  $\mathbf{v}_k = [\mathbf{v}_{k,1}^T \ \mathbf{v}_{k,2}^T \ \cdots \ \mathbf{v}_{k,L}^T]^T$  and  $\mathbf{F}$  is the state transition matrix in Equation (191). Note that this formulation does not require that all  $L$  targets are present at the same time. This enables us to detect a target entering the scene (target birth) and a target leaving the scene (target death). Thus, at any time, many different combinations of target presence, or different modes, are possible. In general, the total number of modes is given by

$$M = \sum_{l=0}^L \frac{L!}{l!(L-l)!} \quad (211)$$

#### 5.5.2.2. State Transition Matrix for PF-TBD

By tracking the state of the mode at each time step, we are inherently estimating when a target is entering or leaving the radar scene. The mode transition is modeled as an  $M$ -state first order Markov chain  $r_k$  and takes values from the set  $\{0, 1, \dots, M-1\}$ . Mode  $r_k = 0$  implies that no targets are present in the scene; mode  $r_k = M-1$  implies that all targets are present in the scene. An  $M \times M$  state transition matrix is constructed based on *a priori* information about each mode transition, assuming that both  $P_B$ , the probability of a new target entering the scene, and  $P_D$ , the probability of a target leaving the scene, are known. The transition matrix for  $L = 2$  is given as

$$\mathbf{\Pi} = \begin{bmatrix} \Pi_{1,1} & P_B(1-P_B) & 0 & 0 \\ P_D(1-P_B) & \Pi_{2,2} & 0 & P_B(1-P_D) \\ P_D(1-P_B) & 0 & \Pi_{3,3} & P_B(1-P_D) \\ P_D^2 & P_D(1-P_D) & P_D(1-P_D) & \Pi_{4,4} \end{bmatrix} \quad (212)$$

with  $\Pi_{1,1} = P_B^2 - P_B + 1$ ,  $\Pi_{4,4} = P_D^2 - P_D + 1$ , and  $\Pi_{2,2} = \Pi_{3,3} = 1 - P_D - P_B + 2P_DP_B$ . We assume that the initial mode probability,  $\Pr(r_0 = m)$ ,  $m = 0, \dots, M-1$ , is known *a priori*.

#### 5.5.2.3. Measurement Model for PF-TBD

For this problem, we assume that the measurements are obtained from video cameras for tracking moving objects. Each measurement frame is assumed to consist of  $A \times B$  cells. The resolution dimension and the center location of the  $(a, b)$ th cell,  $a = 1, 2, \dots, A$ ,  $b = 1, 2, \dots, B$ , are given by  $\rho_a \times \rho_b$  and  $a\rho_a \times b\rho_b$ , respectively. If no targets are present, the measurements in all

the cells are just noise. If a target is present, then the measurements in the cells that are in the vicinity of the target's current position consist of both signal and noise. We consider a point target and a sensor point spread function that can be approximated by a 2-D Gaussian density. Specifically, considering the  $l$ th target,  $l = 1, 2, \dots, L$ , the mode target combination can be formulated as  $M_{r_k, l} = q$ ,  $r_k = 0, 1, \dots, M - 1$ , where the binary valued  $q = 1$  if the  $l$ th target is present in mode  $r_k$ ; otherwise,  $q = 0$ . Based on this, the measurement equation is given by,

$$z_k^{(a,b)} = \begin{cases} \sum_{l=1}^L M_{r_k, l} h_k^{(a,b)}(\mathbf{x}_{k,l}) + w_k^{(a,b)}, & r_k \neq 0 \\ w_k^{(a,b)}, & r_k = 0 \end{cases} \quad (213)$$

where  $h_k^{(a,b)}(x_{k,l}) = \frac{\rho_a \rho_b I_{k,l}}{2\pi\Sigma^2} e^{-E}$ , with  $E = -0.5((a\rho_a - x_{k,l})^2 + (b\rho_b - y_{k,l})^2)/\Sigma^2$  and  $\Sigma$  is a known parameter that controls the video camera blurring. The independent and identically distributed measurement noise samples  $w_k^{a,b}$  are assumed to be zero-mean Gaussian with variance  $r_k$ . All the measurements up to time  $k$  are  $\mathbf{Z}_k = [\mathbf{z}_1 \mathbf{z}_2 \dots \mathbf{z}_k]^T$  and the overall measurement vector is  $\mathbf{z}_k = [z_k^{(1,1)} \dots z_k^{(1,B)} z_k^{(2,1)} \dots z_k^{(2,B)} \dots z_k^{(A,1)} \dots z_k^{(A,B)}]^T$ .

#### 5.5.2.4. Multiple-Mode Multiple-Target TBD Algorithm

The multiple-mode multiple-target TBD algorithm can be formulated as a nonlinear jump Markov process whose kinematic target state vector must take into account the time-varying number of targets at each time step. Given the state and measurement models and the posterior probability density  $p(\mathbf{x}_{k-1}, r_{k-1} | \mathbf{Z}_{k-1})$  at time  $k - 1$ , the task of the multiple-mode multiple-target TBD is to estimate the posterior probability density  $p(\mathbf{x}_k | \mathbf{Z}_k)$  at time  $k$ . This is a multiple-model problem in which the posterior probability density will differ based on the different modes of the target condition. For ease of notation, we define  $r_{k,m}$  to represent  $r_k = m$ , which means that the state mode at time  $k$  is  $m$ . Then, the posterior probability density can be written as,

$$p(\mathbf{x}_k | \mathbf{Z}_k) = \sum_{m=0}^{M-1} p(\mathbf{x}_k, r_{k,m} | \mathbf{Z}_k) \quad (214)$$

where the  $m$  th joint posterior probability density is

$$p(\mathbf{x}_k, r_{k,m} | \mathbf{Z}_k) = p(\mathbf{x}_k | r_{k,m}, \mathbf{Z}_k) \Pr(r_{k,m} | \mathbf{Z}_k) \quad (215)$$

The algorithm steps can be summarized as follows.

- Initialize PDF  $p(\mathbf{x}_0 | r_{0,m}, \mathbf{Z}_{k-1})$  and mode probability  $\Pr(r_{0,m} | Z_0)$ ,  $m = 1, \dots, M - 1$
- For  $k = 1$  through  $K$ , repeat the following steps:
  - Predict the state PDF  $p_{l,m}(\mathbf{x}_k | \mathbf{Z}_{k-1})$ , conditioned on mode  $l$  at time  $k - 1$  and mode  $m$  at time  $k$  for  $l = 0, 1, \dots, M - 1$  and  $m = 1, \dots, M - 1$
  - Compute the likelihood function  $L_{l,m}(\mathbf{z}_k | \mathbf{x}_k)$ , conditioned on modes  $l$  and  $m$
  - Update the posterior state PDF  $p_{l,m}(\mathbf{x}_k | \mathbf{Z}_k)$ , conditioned on modes  $l$  and  $m$

- Compute the mixing probabilities as  $\Pr(r_{k-1,l} | r_{k,m}, \mathbf{Z}_k) = \frac{p_{l,m}(\mathbf{z}_k | \mathbf{Z}_{k-1}) \Pi_{l,m} P_{k-1,l}}{p(\mathbf{z}_k | r_{k,m}, \mathbf{Z}_{k-1}) P_{k-1,m}}$ .
- Using the mixing probabilities, compute the posterior mode probabilities  $\Pr(r_{k,m} | \mathbf{Z}_k)$
- Compute the posterior PDF  $p(\mathbf{x}_k | r_{k,m}, \mathbf{Z}_k) = \sum_{l=0}^{M-1} p_{l,m}(\mathbf{x}_k | \mathbf{Z}_k) \Pr(r_{k-1,l} | r_{k,m}, \mathbf{Z}_k)$  conditioned on a mode.

### 5.5.3. Agile Tracking Using Probability Hypothesis Density Filter

We developed a waveform-agile method for tracking multiple targets as well as estimating the number of targets at each time step using the probability hypothesis density filter (PHDF). The selection process myopically chooses the transmit waveform for each target by minimizing the predicted MSE across a pre-computed waveform library [92]. This work was performed in collaboration with the PI, Muralidhar Rangaswamy.

#### 5.5.3.1. Multiple Target Tracking Model Formulation

We consider a multiple target tracking system that uses  $J$  sensors to track  $L_k$  targets, where  $L_k$  may be known or unknown. The  $i$ th sensor transmits a single Gaussian windowed LFM chirp at time step  $k$ ,  $\tilde{s}_k^{(i)}(t) = \sqrt{2\varepsilon_k} \text{Re}\{s_k^{(i)}(t; \theta_k^{(i)}) e^{j2\pi f_c t}\}$ ,  $i = 1, \dots, J$ , where  $f_c$  is the carrier frequency,  $\varepsilon_k$  is the signal energy, and  $s_k^{(i)}(t; \theta_k^{(i)}) = (\pi(\lambda_k^{(i)})^2)^{-1/4} e^{-0.5t^2/(\lambda_k^{(i)})^2} e^{j2\pi b_k^{(i)} t^2}$  is the transmitted signal. The signal is characterized by the parameters  $\theta_k^{(i)} = [\lambda_k^{(i)} \ b_k^{(i)}]^T$ , where  $\lambda_k^{(i)}$  depends on the waveform duration and  $b_k^{(i)}$  is the signal's FM rate. If we assume that each target generates only one observation at each time step, and that targets generate observations independently of one another, then the  $m$ th (noiseless) received waveform is

$$r_{k,m}^{(i)}(t) = \text{Re}\{A_m^{(i)} \tilde{s}_k^{(i)}(t - \tau_{k,m}^{(i)}) e^{j2\pi(f_c + \nu_{k,m}^{(i)})t}\}, \quad m = 1, \dots, M_k \quad (216)$$

where  $A_m^{(i)}$ ,  $\tau_{k,m}^{(i)}$ , and  $\nu_{k,m}^{(i)}$  are the amplitude fading, time-delay and Doppler-shift changes, respectively, on the transmitted signal. Assuming that the observation originated from a target, we can use the observed time-delay and Doppler shift values, and relationships  $\tau_{k,m}^{(i)} = 2r_{k,m}^{(i)} / c$  and  $\nu_{k,m}^{(i)} = 2f_c \dot{r}_{k,m}^{(i)} / c$ , where  $c$  is the velocity of the waveform in air, to obtain range  $r_{k,m}^{(i)}$  and range-rate  $\dot{r}_{k,m}^{(i)}$  information from the  $i$ th transmitter to that target at time  $k$ . The state vector,  $\mathbf{x}_{k,l} = [x_{k,l} \ \dot{x}_{k,l} \ y_{k,l} \ \dot{y}_{k,l}]^T$ , is in terms of the 2-D Cartesian coordinates of the location  $(x_{k,l}, y_{k,l})$  and velocity  $(\dot{x}_{k,l}, \dot{y}_{k,l})$  of the  $l$ th target,  $l = 1, \dots, L_k$ , at time step  $k$ . The state equation follows the constant velocity model as in Equation (191).

The  $m$ th measurement  $\mathbf{z}_{k,m}^{(i)} = [r_{k,m}^{(i)} \ \dot{r}_{k,m}^{(i)} \ \phi_{k,m}^{(i)}]^T = h^{(i)}(\mathbf{x}_{k,l}) + \mathbf{v}_k^{(i)}$  at the  $i$ th sensor at time  $k$  provides range, range-rate and bearing angle information. Note that the  $m$ th measurement may not necessarily be from the  $l$ th target; it may correspond to a false alarm measurement due to

clutter. Ignoring the noise term,  $\mathbf{z}_{k,m}^{(i)} = h^{(i)}(\mathbf{x}_{k,l})$  depends on  $r_{k,m}^{(i)} = [(x_{k,l} - x_{sr}^{(i)})^2 + (y_{k,l} - y_{sr}^{(i)})^2]^{1/2}$ ,  $\dot{r}_{k,m}^{(i)} = [(x_{k,l} - x_{sr}^{(i)})\dot{x}_{k,l} + (y_{k,l} - y_{sr}^{(i)})\dot{y}_{k,l}] / r_{k,m}^{(i)}$  and  $\phi_{k,m}^{(i)} = \arctan(y_{k,l} / x_{k,l})$ , where  $(x_{sr}^{(i)}, y_{sr}^{(i)})$  is the 2-D (fixed) location coordinates of the  $i$ th sensor. When the  $m$ th observation originates from the  $l$ th target, or  $m = l$ , then the state of the  $l$ th target can be estimated. The measurement noise process  $\mathbf{v}_k^{(i)}$  is assumed zero-mean Gaussian with covariance  $N(\theta_k^{(i)})$  and depends on the  $i$ th transmitted waveform parameters  $\theta_k^{(i)} = [\lambda_k^{(i)} \ b_k^{(i)}]^T$ . Assuming high SNR,  $N_{AF}(\theta_k^{(i)})$  in the block diagonal measurement noise covariance  $N(\theta_k^{(i)}) = \text{diag}(N_{AF}(\theta_k^{(i)}), N_\phi)$  can be approximated by the CRLB on the variance of the range and range-rate estimators, following [78]. The CRLB can be computed using the narrowband AF of the transmitted waveform.  $N_\phi$  does not depend on the waveform parameters; it is set based on the radar scene.

### 5.5.3.2. PHDF and Multiple Target Tracking

The probability hypothesis density filter (PHDF) dynamically estimates the number of targets and their parameters at each time step by modeling them as random finite sets (RFS) [91]. Using RFS, we define the multiple target state RFS as  $\mathbf{X}_k = \{\mathbf{x}_{k,1}, \dots, \mathbf{x}_{k,L_k}\}$ , where  $\mathbf{x}_{k,l}$  is the state vector of the  $l$ th target, the measurement RFS up to time step  $k$  as  $\mathbf{Z}_{1:k} = \{\mathbf{Z}_1, \dots, \mathbf{Z}_k\}$ , and the multiple target measurement RFS as  $\mathbf{Z}_k = \{\mathbf{z}_{k,1}^{(1)}, \dots, \mathbf{z}_{k,M_k}^{(1)}, \mathbf{z}_{k,1}^{(2)}, \dots, \mathbf{z}_{k,M_k}^{(2)}, \dots, \mathbf{z}_{k,1}^{(i)}, \dots, \mathbf{z}_{k,M_k}^{(i)}\}$ , where  $\mathbf{z}_{k,m}^{(i)}$  is the  $m$ th measurement and  $M_k$  is the total number of measurements. If we consider any state vector  $\mathbf{x}_k \in \mathbf{X}_k$ , it can be shown that by integrating the probability hypothesis density or intensity function  $\zeta(\mathbf{x}_k | \mathbf{Z}_{1:k})$  over a given region will provide the expected number of targets in that region; the locations of the density peaks will provide estimates of the target states [91]. The PHDF assumes that the predicted multiple target posterior PDF  $p(\mathbf{x}_k | \mathbf{Z}_{1:k-1})$  can be characterized by the multiple target RFS intensity function  $\zeta(\mathbf{x}_k | \mathbf{Z}_{1:k-1})$ . Thus, given the posterior intensity  $\zeta(\mathbf{x}_{k-1} | \mathbf{Z}_{1:k-1})$  at time step  $(k-1)$ , the predicted intensity  $\zeta(\mathbf{x}_k | \mathbf{Z}_{1:k-1})$  can be obtained as

$$\zeta(\mathbf{x}_k | \mathbf{Z}_{1:k-1}) = \zeta(\mathbf{x}_k^{\text{new}} | \mathbf{Z}_{1:k}) + \int [\text{Pr}_{k|k-1}(\mathbf{x}_{k-1}) p(\mathbf{x}_k | \mathbf{x}_{k-1}) + \zeta(\mathbf{x}_k^{\text{spn}} | \mathbf{Z}_{1:k-1})] \zeta(\mathbf{x}_{k-1} | \mathbf{Z}_{1:k-1}) d\mathbf{x}_{k-1} \quad (217)$$

where  $\mathbf{x}_k^{\text{new}}$  are targets that appear at time step  $k$ ,  $\mathbf{x}_k^{\text{spn}}$  are targets spawning from targets from the previous time  $(k-1)$ , and  $\text{Pr}_{k|k-1}(\mathbf{x}_{k-1})$  is the probability that a target that was present at time  $(k-1)$  will still be present at time step  $k$ . The posterior intensity is given by

$$\zeta(\mathbf{x}_k | \mathbf{Z}_{1:k}) = (1 - \text{Pr}_k^{\text{det}}(\mathbf{x}_k)) \zeta(\mathbf{x}_k | \mathbf{Z}_{1:k-1}) + \sum_{\mathbf{Z}_k \in \mathbf{Z}_{1:k}} \frac{\text{Pr}_k^{\text{det}}(\mathbf{x}_k) p(\mathbf{Z}_k | \mathbf{x}_k) \zeta(\mathbf{x}_k | \mathbf{Z}_{1:k-1})}{\zeta(\mathbf{Z}^{\text{clt}}) + \int \text{Pr}_k^{\text{det}}(\tilde{\mathbf{x}}_k) p(\mathbf{Z}_k | \tilde{\mathbf{x}}_k) \zeta(\tilde{\mathbf{x}}_k | \mathbf{Z}_{1:k-1}) d\tilde{\mathbf{x}}_k} \quad (218)$$

where  $\text{Pr}_k^{\text{det}}(\mathbf{x}_k)$  is the probability of detecting a target at time step  $k$ . We assume that the measurements can include possible false alarms,  $\mathbf{Z}^{\text{clt}}$ , with clutter intensity  $\zeta(\mathbf{Z}^{\text{clt}})$ . The clutter is assumed independent from the target measurements.

The PHDF prediction and update equations in Equations (217) and (218), respectively, involve multiple integrals that are not computationally tractable. Following a particle filter (PF) implementation (PF-PHDF) [95] instead, the intensity function  $\zeta(\mathbf{x}_{k-1}|\mathbf{Z}_{1:k-1})$  at time  $(k-1)$  is approximated using particles  $\mathbf{x}_{k-1}^{(n)}$  and their corresponding weights  $w_{k-1}^{(n)}$ ,  $n = 1, \dots, N$ , as

$$\zeta(\mathbf{x}_{k-1}|\mathbf{Z}_{1:k-1}) = \sum_{n=1}^N w_{k-1}^{(n)} \delta(\mathbf{x}_{k-1} - \mathbf{x}_{k-1}^{(n)}) \quad (219)$$

where  $N$  is the number of particles and  $\delta(\cdot)$  is the Dirac delta function. If we substitute Equation (219) into (217), then

$$\zeta(\mathbf{x}_k|\mathbf{Z}_{1:k-1}) = \zeta(\mathbf{x}_k^{\text{new}}|\mathbf{Z}_{1:k}) + \sum_{n=1}^N w_{k-1}^{(n)} [\text{Pr}_{k|k-1}(\mathbf{x}_{k-1}^{(n)}) p(\mathbf{x}_k|\mathbf{x}_{k-1}^{(n)}) + \zeta(\mathbf{x}_k^{\text{spn}}|\mathbf{Z}_{1:k-1})] \quad (220)$$

A particle approximation of  $\zeta(\mathbf{x}_k|\mathbf{Z}_{1:k-1})$  is obtained using importance sampling on each term in Equation (220). In particular,  $\mathbf{x}_k^{(n)} = \mathbf{F}\mathbf{x}_{k-1}^{(n)} + \mathbf{w}_{k-1}$ ,  $n = 1, \dots, N$ , and  $\mathbf{x}_k^{(n)}$ ,  $n = N+1, \dots, N+N_\xi$  are drawn uniformly from the state space that the targets will occupy. Then, (220) can be approximated by

$$\zeta(\mathbf{x}_k|\mathbf{Z}_{1:k-1}) = \sum_{n=1}^{N+N_\xi} w_{k|k-1}^{(n)} \delta(\mathbf{x}_k - \mathbf{x}_k^{(n)}) \quad (221)$$

where  $N_\xi$  is the number of particles for the new targets at time  $k$ ,  $w_{k|k-1}^{(n)} = \text{Pr}_{k|k-1}(\mathbf{x}_{k-1}^{(n)}) w_{k-1}^{(n)}$  for  $n = 1, \dots, N$  and  $w_{k|k-1}^{(n)} = 1/N_\xi$  for  $n = N+1, \dots, (N+N_\xi)$ . In our simulations, we assumed that the transitional prior intensity  $\zeta(\mathbf{x}_k^{\text{new}}|\mathbf{Z}_{1:k}) = 0$  and  $\zeta(\mathbf{x}_k^{\text{spn}}|\mathbf{Z}_{1:k-1})$  is a uniform distribution with the bound of the state space set *a priori*. We can thus obtain a particle approximation of the posterior intensity function  $\zeta(\mathbf{x}_k|\mathbf{Z}_{1:k})$  as

$$\zeta(\mathbf{x}_k|\mathbf{Z}_{1:k}) = \sum_{n=1}^{N+N_\xi} w_k^{(n)} \delta(\mathbf{x}_k - \mathbf{x}_k^{(n)}) \quad (222)$$

where

$$w_k^{(n)} = \left( 1 - \text{Pr}_k^{\text{det}}(\mathbf{x}_k^{(n)}) + \sum_{\mathbf{Z}_k \in \mathbf{Z}_{1:k}} \frac{\text{Pr}_k^{\text{det}}(\mathbf{x}_k^{(n)}) p(\mathbf{Z}_k|\mathbf{x}_k^{(n)})}{\zeta(\mathbf{Z}_k^{\text{clt}}) + C_k(\mathbf{Z}_k)} \right) w_{k-1|k}^{(n)} \quad (223)$$

$$C_k(\mathbf{Z}_k) = \sum_{n=1}^{N+N_\xi} \text{Pr}_k^{\text{det}}(\mathbf{x}_k^{(n)}) p(\mathbf{Z}_k|\mathbf{x}_k^{(n)}) w_{k|k-1}^{(n)}$$

The particles and their weights are divided into  $L_k$  groups using a clustering algorithm such as k-means. The number of clusters  $L_k$  is either assumed known or estimated directly from  $\zeta(\mathbf{x}_k | \mathbf{Z}_{1:k})$ . Denoting the particles and weights from the  $l$ th target,  $l = 1, \dots, L_k$  by  $\mathbf{x}_{k,l}^{(n)}$ , and  $w_{k,l}^{(n)}$ , respectively, then the  $l$ th target state is estimated as  $\hat{\mathbf{x}}_{k|k,l} = \sum_{l \in \ell} w_{k,l}^{(\ell)} \mathbf{x}_{k,l}^{(\ell)}$ ; its covariance is given by  $\mathbf{P}_{k|k,l} = \sum_{l \in \ell} w_{k,l}^{(\ell)} (\mathbf{x}_{k,l}^{(\ell)} - \hat{\mathbf{x}}_{k|k,l})(\mathbf{x}_{k,l}^{(\ell)} - \hat{\mathbf{x}}_{k|k,l})^T$ .

### 5.5.3.3. Waveform Optimization

The optimal transmit waveform is the one that minimizes the target tracking error, which is approximated by the predicted MSE. For  $\mathbf{x}_k \in \mathbf{X}_k$  and estimated by with  $\hat{\mathbf{x}}_k$  given  $\mathbf{Z}_{1:k}$ , this error is given by  $J(\theta_k) = E_{\mathbf{x}_k, \mathbf{Z}_k | \mathbf{Z}_{1:k-1}} \left( (\mathbf{x}_k - \hat{\mathbf{x}}_k)^T \Lambda (\mathbf{x}_k - \hat{\mathbf{x}}_k) \right)$ , where  $\Lambda$  is a weighting matrix that sets the relative importance of position and velocity errors. Although the minimization of  $J(\theta_k)$  can be performed using algorithms such as the gradient descent, these algorithms are computationally intensive to be used effectively. Here, we use the unscented transform to approximate  $J(\theta_k)$ , where the first and second moments of the underlying distribution are estimated using a small set of deterministically chosen particles or sigma points.

From the PF-PHDF, the state estimate of the  $l$ th target at time step  $(k-1)$  is  $\hat{\mathbf{x}}_{k-1|k-1,l}$ ; its corresponding error covariance matrix is  $\mathbf{P}_{k-1|k-1,l}$ . The state equation can then be used to propagate the target state and covariance matrix forward in time using  $\hat{\mathbf{x}}_{k|k-1,l} = \mathbf{F} \hat{\mathbf{x}}_{k-1|k-1,l}$  and  $\mathbf{P}_{k|k-1,l} = \mathbf{F} \mathbf{P}_{k-1|k-1,l} \mathbf{F}^T + \mathbf{Q}$ . Then  $\mathbf{P}_{k|k,l}$  is obtained using the unscented transform and  $\mathbf{P}_{k|k-1,l}$ , using using a set of  $2S+1$  sigma points  $\chi_{k,j,l}$  and corresponding weights  $w_{k,j,l}$ ,  $j = 1, \dots, (2S+1)$ . Specifically, the observation state,  $\mathbf{z}_{k,l}^{(i)} | \mathbf{Z}_{1:k-1,l}$ , of the  $i$ th sensor is approximated into the measurement space using  $\mathbf{z}_{k,j,l}^{(i)} = h^{(i)}(\chi_{k,j,l})$ ,  $j = 1, \dots, (2S+1)$ . Note that, while there is a set of state sigma points  $\chi_{k,j,l}$  for the  $l$ th target, the measurement sigma points  $\mathbf{z}_{k,j,l}^{(i)}$  are calculated for each sensor that observes the  $l$ th target. If there are  $L_k$  sets of state sigma points, there will be  $L_k \times J$  sets of measurement sigma points. The unscented transform uses the sigma points to update the covariance matrices

$$\mathbf{P}_{ZZ}^{(i)} = \sum_{j=0}^{2S} w_{k,j,l} (\mathbf{z}_{k,j,l}^{(i)} - \bar{\mathbf{z}}_{k,l}^{(i)}) (\mathbf{z}_{k,j,l}^{(i)} - \bar{\mathbf{z}}_{k,l}^{(i)})^T \quad (224)$$

$$\mathbf{P}_{XZ}^{(i)} = \sum_{j=0}^{2S} w_{k,j,l} (\chi_{k,j,l} - \bar{\chi}_{k,l}) (\mathbf{z}_{k,j,l}^{(i)} - \bar{\mathbf{z}}_{k,l}^{(i)})^T \quad (225)$$

with

$$\bar{\chi}_{k,l} = \sum_{j=0}^{2S} w_{k,j,l} \chi_{k,j,l} \quad (226)$$

and

$$\bar{\mathbf{Z}}_{k,l}^{(i)} = \sum_{j=0}^{2S} W_{k,j,l} \mathbf{Z}_{k,j,l}^{(i)} \quad (227)$$

These covariances, together with  $\mathbf{P}_{k|k-1,l}$  are then used to update the posterior covariance as a function of the state as  $\mathbf{P}_{k|k,l}^{(i)}(\theta_k^{(i)}) = \mathbf{P}_{k|k-1,l} - \mathbf{P}_{k|k}^{\text{upd}}$  and  $\mathbf{P}_{k|k}^{\text{upd}} = \mathbf{P}_{XZ}^{(i)}[\mathbf{P}_{ZZ}^{(i)} + N(\theta_k^{(i)})]^{-1}(\mathbf{P}_{XZ}^{(i)})^T$ . Note that these equations need to be applied for each sensor in order to configure the transmit waveform of each sensor. They are first applied with  $i=1$ . Then, for  $i=2, \dots, J$ , they are repeated, but with  $\mathbf{P}_{k|k-1,l}$  replaced by  $\mathbf{P}_{k|k,l}^{(i-1)}(\theta_k^{(i-1)})$ . The final predicted covariance is  $\mathbf{P}_{k|k,l}^{(J)}(\theta_k^{(J)})$ . After updating the covariances of all sensors, and with  $\Theta_k = [\theta_k^{(1)} \dots \theta_k^{(J)}]^T$ , the predicted MSE for a single target can be approximated as  $J(\Theta_k) \approx \text{trace}(\Lambda \mathbf{P}_{k|k,l}(\Theta_k))$ . The waveform selection algorithm chooses the set of transmit waveforms for  $J$  sensors that minimize the predicted MSE, that is,  $\hat{\Theta}_k = \arg \min_{\Theta_k} \{ \sum_{l=1}^{L_k} J_l(\Theta_k) \}$ . A library of waveforms, with varying parameters  $\theta$  ranging over  $[\lambda_{\min}, \lambda_{\max}]$  and  $[\text{BW}_{\min}, \text{BW}_{\max}]$ , can be determined prior to tracking to fit the characteristics of a radar system.

## 5.6. Simulation and Experimental Results

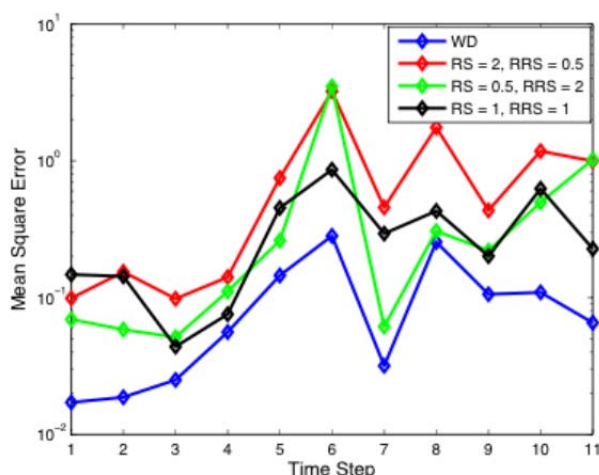
### 5.6.1. Waveform-Agile Multi-Modal TBD

We consider a single target moving in a random trajectory, starting at  $(-13, -39.98)$  ft with initial velocity of  $(1.1891, 2.3266)$  ft/s at 8 dB SNR. The PF-TBD algorithm used 1,500 particles and 20 iterations were used. A dictionary of 25 LFM chirp signals was used, formed by varying range resolution from 0.5 to 2 ft and range-rate resolution from 0.5 to 2 ft/s; the resulting LFM parameters are provided in Table 3

**Table 3. Parameters of LFM Chirp Library**

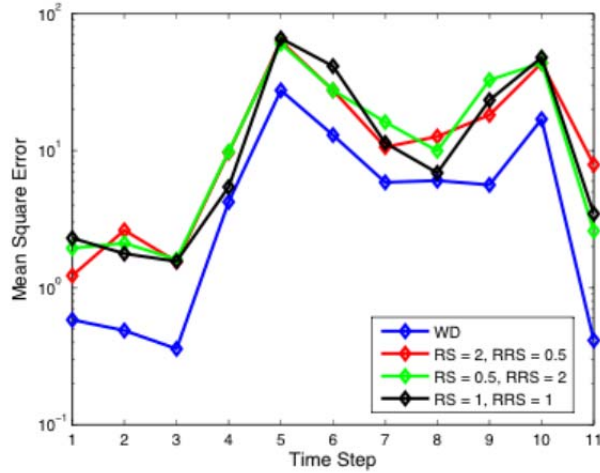
	Range-Rate Resolution (ft/s)				
	0.5	0.707	1	1.414	2
0.5 ft range resolution					
Duration (ms)	0.164	0.116	0.082	0.058	0.041
FM rate ( $10^{12}$ Hz <sup>2</sup> )	6.283	8.886	12.566	17.772	25.133
0.707 ft range resolution					
Duration (ms)	0.164	0.116	0.082	0.058	0.041
FM rate ( $10^{12}$ Hz <sup>2</sup> )	4.443	6.283	8.886	12.566	12.772
1 ft range resolution					
Duration (ms)	0.164	0.116	0.082	0.058	0.041
FM rate ( $10^{12}$ Hz <sup>2</sup> )	3.142	4.443	6.283	8.886	12.566
1.414 ft range resolution					
Duration (ms)	0.164	0.116	0.082	0.058	0.041
FM rate ( $10^{12}$ Hz <sup>2</sup> )	2.221	3.142	4.443	6.283	8.886
2 ft range resolution					
Duration (ms)	0.164	0.116	0.082	0.058	0.041
FM rate ( $10^{12}$ Hz <sup>2</sup> )	1.571	2.221	3.142	4.443	6.283

MSE comparison plots are shown in Figure 33, Figure 34, and Figure 35 using the PF-TBD algorithm with waveform selection and fixed waveforms for different weight vectors  $\mathbf{g} = [g_x, g_{\dot{x}}, g_y, g_{\dot{y}}]^T$ . As it can be seen, the tracking performance improves with waveform design.

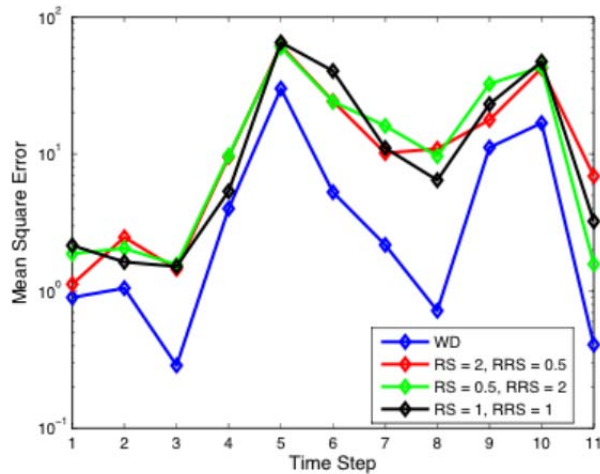


**Figure 33. MSE Comparison of PF-TBD Estimation Position Error, Weight = (1,1,0,0)  
Fixed Waveforms and Waveform Design (WD) for a Random Trajectory**





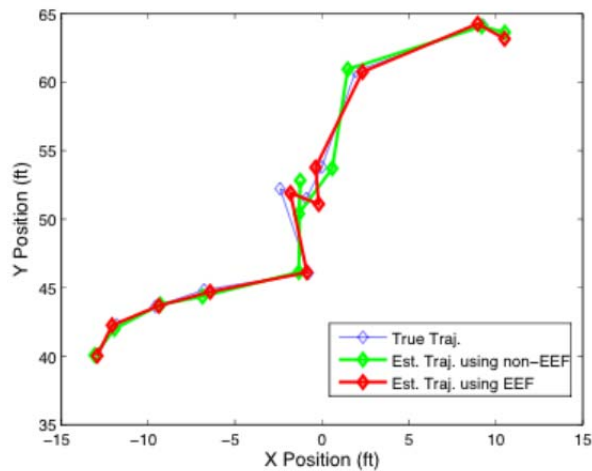
**Figure 34. MSE Comparison of PF-TBD Estimation Position Error, Weight = (1,1,1,1) Fixed Waveforms and Waveform Design (WD) for a Random Trajectory**



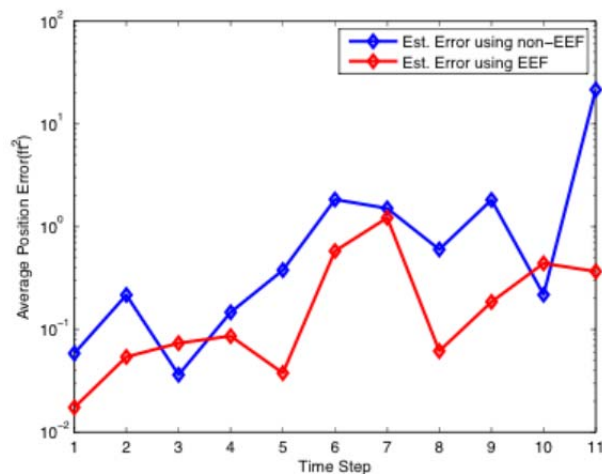
**Figure 35. MSE Comparison of PF-TBD Estimation Position Error, Weight = (0,0,1,1) Fixed Waveforms and Waveform Design (WD) for a Random Trajectory**

### 5.6.2. Dependent Measurements

We compared the tracking performance of the PF-TBD algorithm integrated with the EEF approach with the performance of the PF-TBD assuming that the RF-EO measurements are independent. The same random trajectory and running parameters were used, similar to Section 5.6.1. The simulation results in Figure 36 and Figure 37 demonstrate that the EEF approach results in a lower MSE than when the measurements are assumed independent.



**Figure 36. True and Estimated Trajectory Using PF-TBD Dependent (EEF) and Independent (non-EEF) RF-EO measurements**



**Figure 37. Averaged Position Error Using PF-TBD EEF and Independent Measurements (non-EEF)**

### 5.6.3. Multi-Modal Sensing Platform and Experimental Results

The performance improvement proposed by multi-modal sensing with waveform agility was tested using our team's experimental platform. The platform, designed by Co-PI Ram Narayanan at Pennsylvania State University, consists of a Tektronix AWG710 waveform generator, Agilent MSO8104A mixed signal oscilloscope and Sony cybershot DSC-W30 6 Mpixels. The arbitrary waveform generator (AWG) enables generation of LFM chirp waveforms of different bandwidths. The AWG operates at a maximum rate of 4 GSa/s, making it possible to generate waveforms of frequencies up to 2 GHz. The test-bed platform also includes amplifiers, transmitting and receiving antennas, and a high-speed oscilloscope for recording the received signal. The equipment is shown in Figure 38. The power amplifier has a gain of around 30 dB

and the maximum power output is 28 dBm. The low noise amplifier has a typical gain of 25 dB and noise level of 1.5 dB at 1 GHz. The mixer has a conversion loss of 6.5 dB at operating conditions. The frequency spectrum of the transmitted signal lies within the 1,000 to 1,640 MHz band. Frequency translation is performed to down convert the received signal from the 1,000-1,640 MHz band to the 300-940 MHz range. The oscilloscope samples and records the return signal at 4 GSa/s. The return is processed by software which decides whether further processing is required and what bandwidth waveform must be used at the next time step, if any. A workstation with a GPIB controller enables the software to control the AWG and the oscilloscope. Hence, the radar is capable of automatically make decisions about additional processing and required bandwidth. The radar system parameters are shown in Table 4. Tektronix ArbExpress is used to generate rectangular envelope LFM chirp signals with 1 GHz carrier frequency. The parameters of the library of the LFM signals we used in the experiments is summarized in Table 5.



**Figure 38. RF-EO Experimental Sensing Platform at Pennsylvania State University**

**Table 4. Platform Radar System Parameters**

Parameter	Value
Radar waveform	LFM pulse
Radar bandwidth	40-640 MHz
Pulse width	16 $\mu$ s
Transmit power	$\sim 0.5$ W
Maximum radar range	$\sim 25$ m

**Table 5. Parameters of Waveform Library**

Amplitude, V	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
Duration, ms	16	16	16	16	16	16	16	16	16	16
FM rate $10^{12}$ Hz <sup>2</sup>	3.75	7.5	11.25	15	18.75	22.5	26.25	30	33.75	37.5
Range Res. ft	8.587	4.294	2.862	2.147	1.717	1.431	1.227	1.073	0.954	0.859

The left and right view of the platform are shown in Figure 39 and Figure 40, respectively. The vertical view of the scene is captured by a camera at each time step to obtain angle information. The field is a trapezoid, as shown in Figure 41 (vertical view) and Figure 42 (horizontal view). Also shown are the locations of the platform and the calibrations; these calibrations are used to measure angle information from pictures.

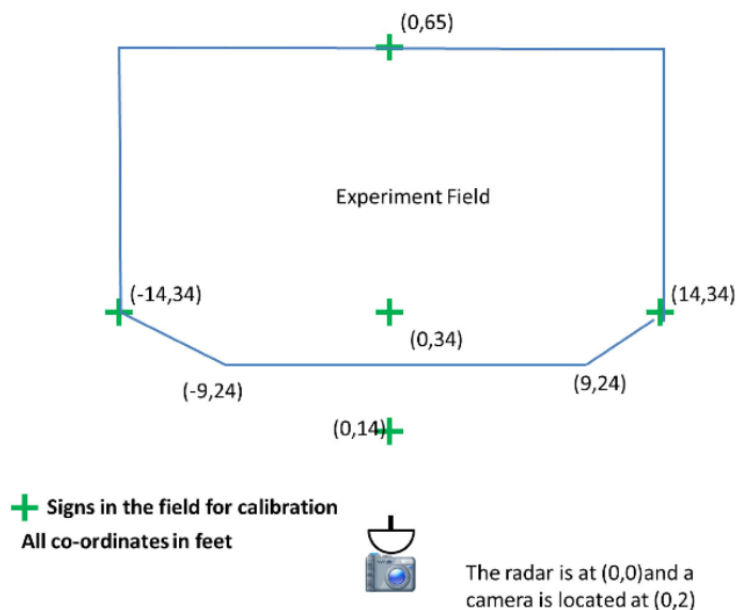


**Figure 39. Experiment Field View: Left View of the Platform**

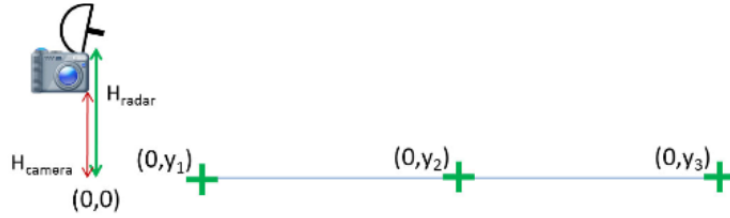


**Figure 40. Experiment Field View: Right View of the Platform**

Using this experimental platform, we demonstrated the advantage of waveform agility in multi-modal tracking using a 2x2 ft trihedral corner reflector as the target. Using the random trajectory used in Sections 5.6.1 and 5.6.2, we compared the estimation position error of the PF-TBD algorithm and the PF-TBD integrated with EEF. The average position error is defined as the mean of position error at 10 time steps, which evaluates the overall tracking ability. The results summarized in Table 6 showed that integrating the EEF into the PF-TBD resulted in reducing the estimation error.



**Figure 41. Experiment Field Layout: Vertical View.**



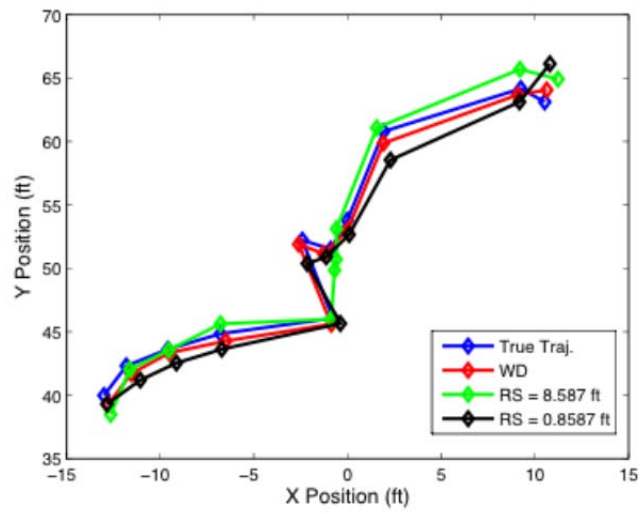
**Figure 42. Experimental Field Layout: Horizontal view**  
(See calibration coordinates in Figure 41)

**Table 6. Averaged Estimation Position Error**

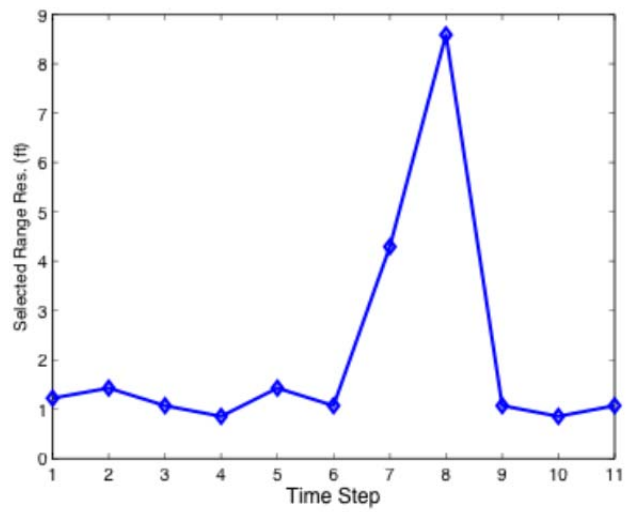
FM rate $10^{12} \text{ Hz}^2$	Random Trajectory	
	MSE using EEF	MSE without using EEF
3.75	1.28	1.43
7.5	1.13	1.87
11.25	1.02	1.99
15	2.29	2.61
18.75	1.78	1.82
22.5	2.73	2.57
26.25	2.01	2.96
30	3.27	3.36
33.75	3.68	4.22
37.5	3.29	3.54

We also demonstrated the advantage of waveform design using the random trajectory and allowing different waveforms to be transmitted at different times. Figure 43 shows the true and estimated random trajectory, and Figure 44 shows that the platform adaptively configures the sensor resolutions to minimize estimation position error. The actual estimation error performance, compared to fixed range resolution waveforms, is shown in Figure 45.

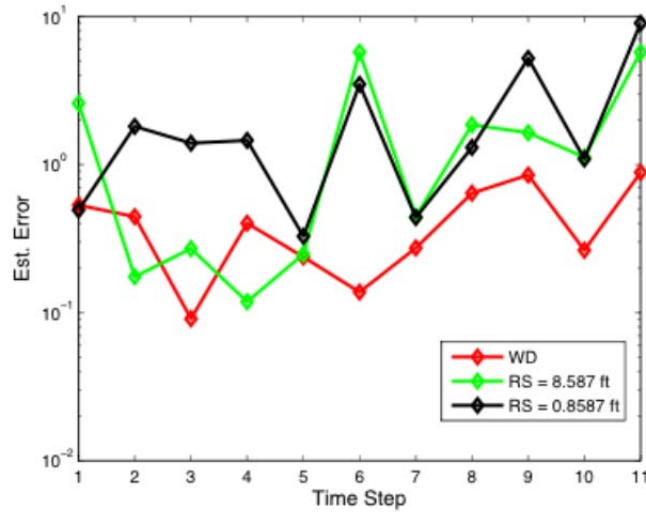




**Figure 43. True and Estimated Random Trajectory from Experimental Setup.**



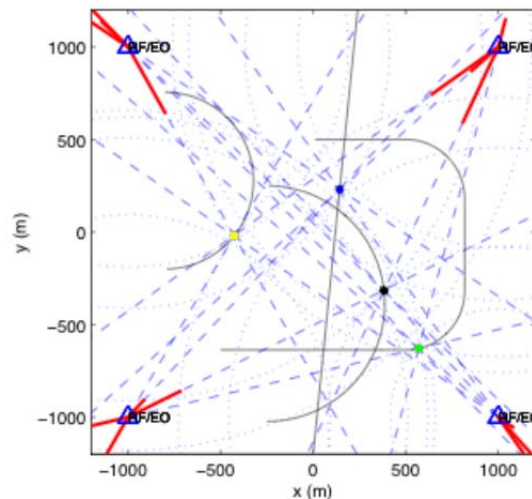
**Figure 44. Optimal Sensor Configuration.**



**Figure 45. Position MSE from Experimental Results Demonstrating Waveform Design for PF-TBD**

#### 5.6.4. Multi-Modal RF-EO Sensing of Multiple Targets

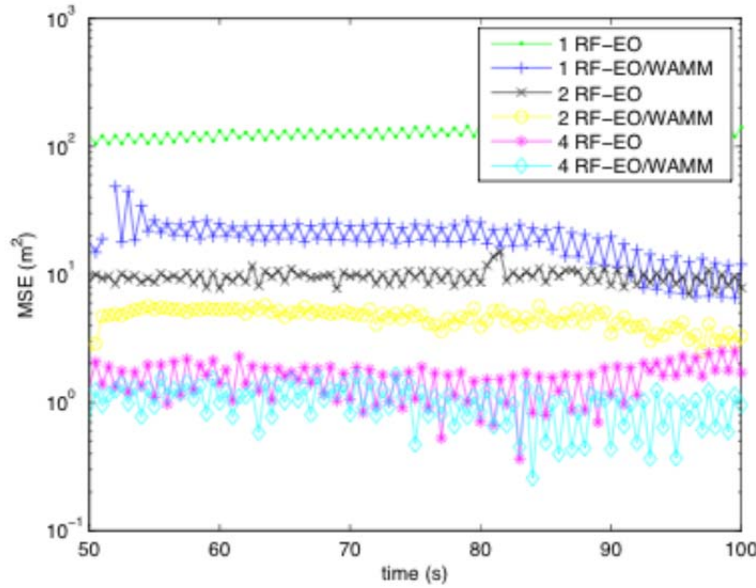
We demonstrated the joint RF-EO multiple target tracking scenario in Figure 46 using 4 RF-EO sensors to alternatively provide measurements to the data processing center, i.e., the RF sensors provided range and range-rate measurements at odd times whereas the EO sensors provided angle measurements at even times. Targets 1, 2, 3, and 4 enter the scene at time steps 1, 50, 100, and 150, respectively. The targets are tracked using range measurements (dotted arcs), range-rate measurements (lines), and angle measurements (dotted lines). As we can see, the trajectory begins with random selected positions (denoted by circles) and then follow the true trajectory.



**Figure 46. Joint RF-EO Multi-Target Tracking, 4 Separate Targets in 2-D space**  
**Target 1 at (559.7,-630.9) m (green), Target 2 at (114.0,240.0) m (blue), Target 3 at (383.8,-326.7) m (black), and Target 4 at (-419.7,-11.2) m (yellow).**



Figure 47 shows the tracking MSE of the position estimate of the first target using different numbers of RF-EO sensors, with and without waveform agility. As we can see, the best performance is obtained when 4 RF-EO sensors are used with waveform selection and adaptive control. Thus, waveform design improves tracking performance, and it substantially improves it over the performance of a single RF-EO sensor tracking case.

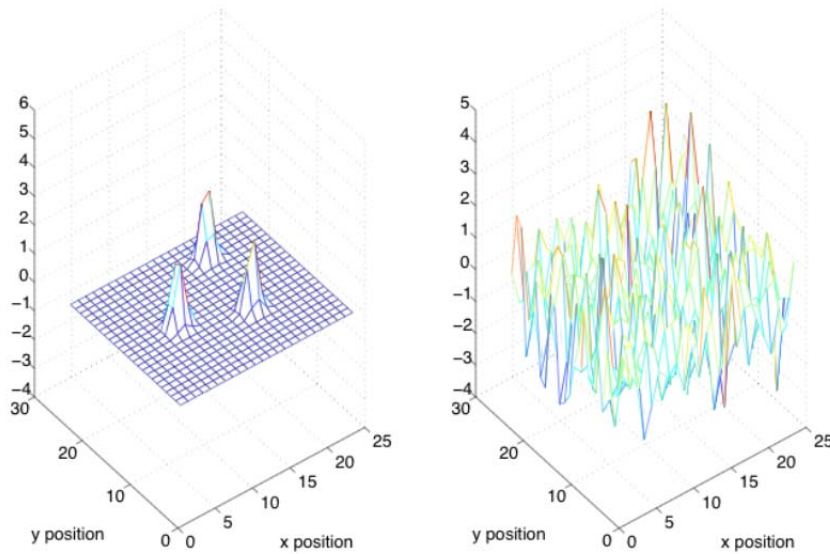


**Figure 47. MSE (Target 1) for MTT with an Increasing Number of RF-EO Sensors**

#### 5.6.5. Multiple Target TBD

We implemented the multiple target PF-TBD algorithm to track three low SNR targets in 2-D. We assumed constant velocity target motion and additive Gaussian process noise models. The state model  $5 \times 5$   $\mathbf{F}$  matrix has zero elements except it is unity along the diagonal, and  $\mathbf{F}_{34} = \mathbf{F}_{12} = \delta t$ , where  $\delta t$  is the duration between time steps. The  $5 \times 5$  covariance matrix  $\mathbf{Q}$  for the modeling error process  $v_k$  is zero except for elements  $\mathbf{Q}_{11} = \mathbf{Q}_{33} = q_1 \delta t^4/4$ ,  $\mathbf{Q}_{12} = \mathbf{Q}_{21} = q_1 \delta t^3/2$ ,  $\mathbf{Q}_{43} = \mathbf{Q}_{34} = q_1 \delta t^3/2$ ,  $\mathbf{Q}_{22} = \mathbf{Q}_{44} = q_1 \delta t^2$ ,  $\mathbf{Q}_{55} = q_2 \delta t$ , where  $q_1$  and  $q_2$  are modeling error process parameters for the target motion and intensity, respectively. For the three targets, the measurements were generated such that the first target enters during frame 4 and leaves during frame 20; the second target enters during frame 10 and leaves during frame 26; and the third target enters during frame 16 and leaves during frame 32. The initial positions and velocities for each of the targets were (7.2, 7.2) m and (0.75, 0.55) m/s, (4.2, 16.2) m and (0.75, -0.75) m/s, and (9, 20) m and (0.05, -0.85) m/s, respectively.

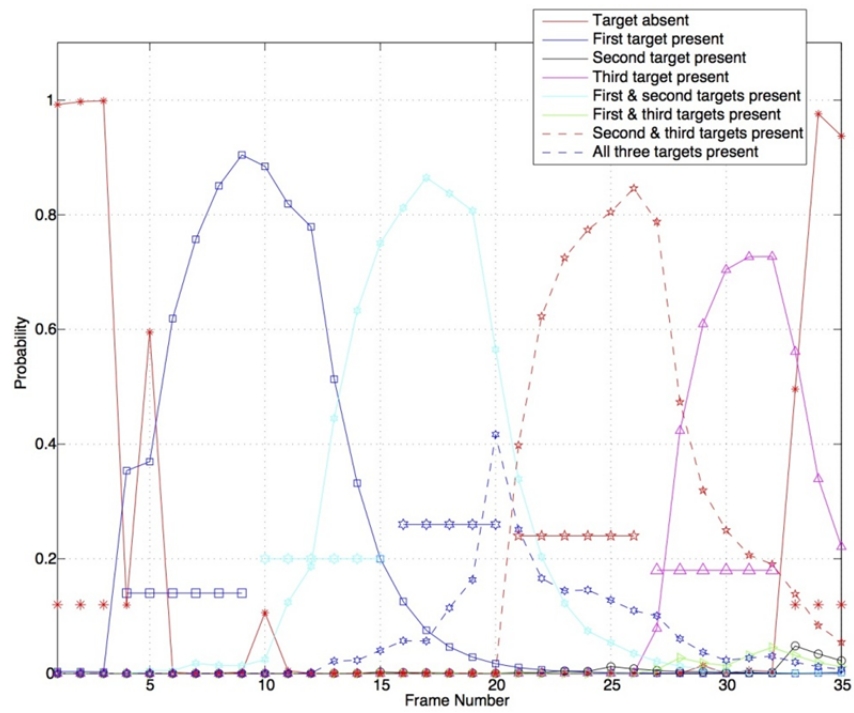
Figure 48 shows frame 20 with and without 9 dB peak SNR noise.



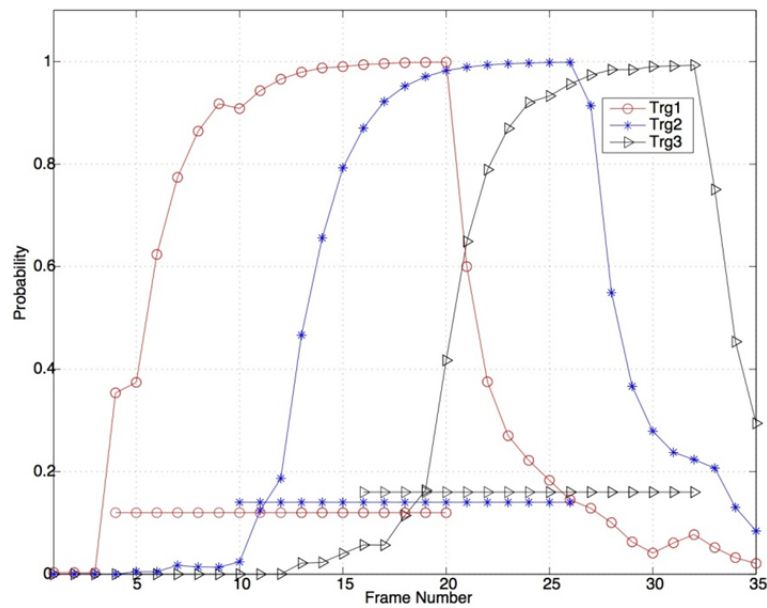
**Figure 48. Measurement with Three Targets  
Signal Only (left) and 9 dB Peak SNR Noisy Signal (right).**

The process noise parameters are set to  $q_1 = 0.001$  and  $q_2 = 0.01$ , and the measurement noise variance is 1. The probabilities of entering and leaving the scene are chosen as  $P_B = P_D = 0.05$ . The parameters used in generating new particles are a measurement threshold of  $\gamma = 2.25$ , target velocity uniform range values from -1 to 1 m/s, and target intensity range values from 10 to 30. The expected peak SNR ranged from 0 to 20 dB.

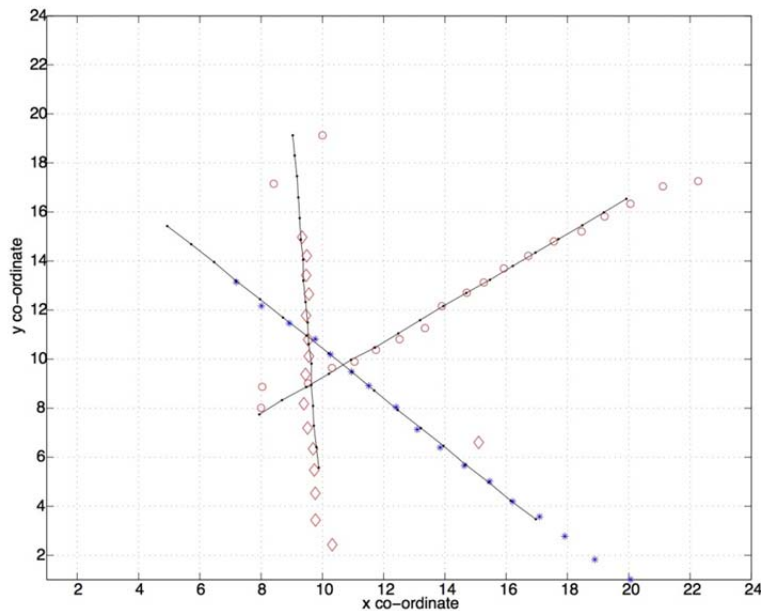
Figure 49 shows the estimated mode probability at different frames for 9 dB peak SNR. The plots at the bottom of the figure (around  $y = 0.2$  m) show the true state of the mode. As it can be seen, the proposed algorithm closely follows the true mode transition. Figure 50 shows the probability of target presence for all three targets. The bottom part of the figure shows the true target presence at different frames. Figure 51 shows the true and estimated target position; it can be seen that the proposed algorithm can closely track all three targets at low SNR.



**Figure 49. Mode Probability for 3 Targets at 9 dB Peak SNR**



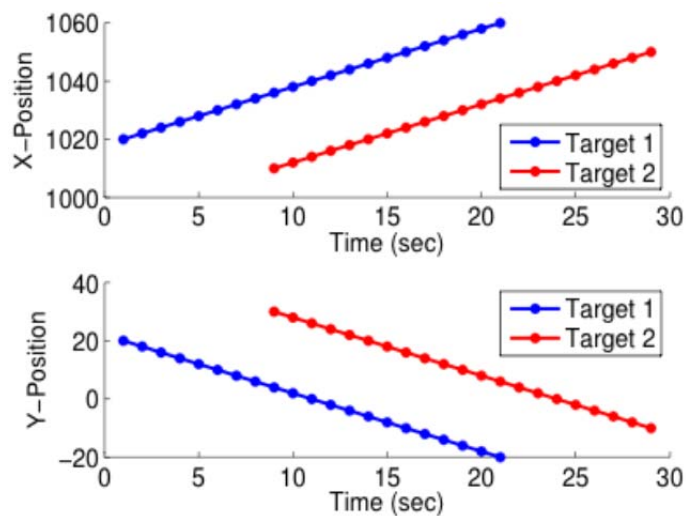
**Figure 50. Target Presence Probability for 3 Targets at 9 dB Peak SNR**



**Figure 51. True and Estimated Target Trajectories of 3 Targets at 9 dB Peak SNR.**

#### 5.6.6. PHDF and Multiple Target Tracking

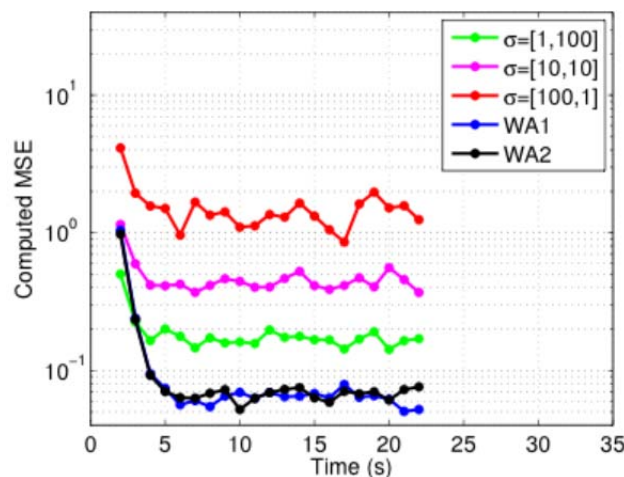
In a two-radar tracking scenario, we used the PF-PHDF to track two targets at each time step. The radars are located at (0, 0) km and (1, 1) km, while the target tracks are plotted in Figure 52. In the first 8 s of the tracking scenario, Target 1 is observed alone, then both targets are present for the next 13 s, and then Target 2 is observed alone for the last 8 s of the tracking scenario.



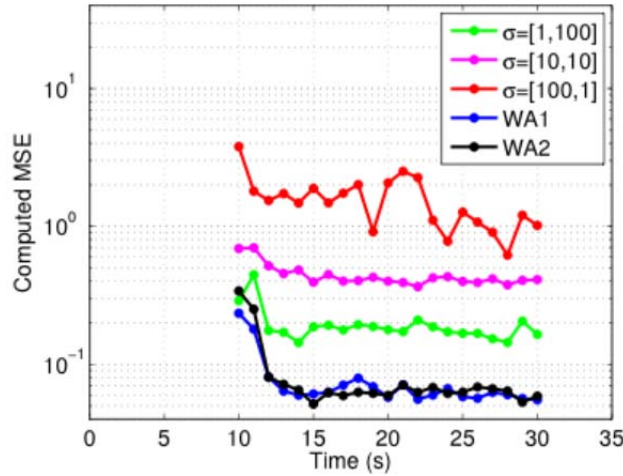
**Figure 52. Target Trajectories for Simulated Target 1 and Target 2**

At each time step, the transmit waveforms for each radar are chosen based on the minimum predicted MSE. Figure 53 and Figure 54 show the MSE from transmitting fixed waveforms, as well as from transmitting dynamically selected waveforms; for dynamic selection, the MSE is provided for the case of both radars transmitting the same waveform as well as each radar transmitting a different waveform. In the computed MSE, position errors are given five times the weighting of velocity errors so that the algorithm chooses waveforms with good range resolution.

Dynamically selecting the transmitted waveforms reduced the target tracking error over a range of fixed waveforms. Although not true in general, for this specific tracking geometry there is little benefit to transmitting different waveforms because the optimal waveforms are approximately the same. In scenarios with more complex radar and target geometries, as well as more complex target tracks, this will not be the case. Additionally, the tracking error for Targets 1 and 2 appear to be independent. While this is not true in general, it is in this scenario because both targets are located in the same area relative to the radars, and the waveform choices are optimal waveform choices are approximately the same for each target.



**Figure 53. Computed MSE for Target 1 for Same and Different Transmit Waveforms**  
**WA1 is for the MSE when both radars transmit the same waveform; WA2 for the MSE**  
**when each radar transmit a different waveform**



**Figure 54. Computed MSE for Target 2**

## 5.7. Conclusions

The track-before-detect algorithm, implemented sequentially using particle filtering (PF-TBD), was applied to the asymmetric RF-EO multi-modal platform to track low observable targets. The algorithm was integrated with waveform agile selection to deal with targets with different energy returns and to reduce information loss due to low SNR or when dependent measurements from different sensors are not appropriately processed. The PF-TBD was also integrated with the embedded exponential family approach in order to take into consideration the statistical dependence between RF and EO measurements when used at the same time to observe the same scene. The improvements in tracking performance were demonstrated not only with simulations but also using an experimental platform, consisting of both RF and ED sensors.

The PF-TBD was extended to estimating the parameters of multiple targets in low SNR. We first considered joint multi-modal sensing based on using dynamic agility selection to optimize the tracking performance of multiple maneuvering targets. The algorithm used waveform design for RF sensing and resolution switching modes for EO sensing when both sensor measurements experienced high false alarm rates. It also made use of the Rao-Blackwellized particle filter to track an unknown number of targets within the adaptive configuration framework. Secondly, we considered a new algorithm that estimates the joint posterior probability density of all the target trajectories while keeping track of targets entering and leaving the noisy radar scene under observation by considering the different scenarios as multiple modes. Also for multiple target tracking, we also considered the probability hypothesis density filter, implemented using particle filtering and integrated with waveform agility. The selection process myopically chooses the transmit waveform for each target by minimizing the predicted MSE across a pre-computed waveform library.

## **6. Radar Test-Bed Development**

### **6.1. Methods, Assumptions, and Procedures**

Increasingly complex target scenarios call for sophisticated techniques such as waveform and sensing diversity for resolving individual targets or target scattering centers for target identification and recognition. Waveform design is therefore an essential ingredient of modern radar systems [96]. With the available electromagnetic (EM) spectrum becoming increasingly scarce, a crucial requirement is one of multimodal sensor operation with fully adaptive waveform capability [97], [98].

Multifunction radio frequency (RF) systems have been studied for a long time [99]–[101]. A scalable multifunction RF system allows the RF functionality (radar, electronic warfare and communications) to be easily extended and the RF performance to be scaled to the requirements of different missions and platforms. Such a system, explained in [99], has a design that is functionally partitioned into scalable, reusable modules consisting of isolated self-contained functional elements. The architecture consists of multifunction apertures that are active phased array antennas that can be changed in size, shifted in RF and switched in polarization. Technology advances such as analog-to-digital conversion speed, reducing hardware size have allowed an increasing degree of digitization, and hence modularity and flexibility. The Advanced Multifunction RF System (AMRFS) initiated by the Office of Naval Research was a demonstration of the concept of common broadband apertures being used to simultaneously perform a large number of RF functions [100]. The underlying concept of AMRFS is to divide the frequency band into an optimal number of segments based on cost and functionality and then utilize separate, electronically scanned, solid state transmit and receive apertures. The test-bed consists of four low band and high band transmit and receive array apertures. Fundamental to the selected AMRFS architecture is the physical separation of receive and transmit and the partitioning of the 1–18 GHz frequency coverage into two bands, low band (1–5 GHz) and high band (4–18 GHz).

Multifunction radar systems have also been developed [102]–[105]. These are phased array systems that can perform a wide variety of radar functions such as tracking, surveillance, etc. [106]. MESAR (Multi-function electronically scanned, adaptive radar) is one such system [103]. It provided the functions of surveillance, rapid track initiation and multiple target tracking using an array of 918 elements. By the use of multiple, stepped frequency narrow bandwidth waveforms over the agile bandwidth of the radar, and processing the returned signals to synthesize a high bandwidth response, a range resolution of less than 1 m was obtained. It also supported adaptive tracking by having the update rates and waveforms used by an adaptive function of the target environment. The APAR multifunction radar used four fixed active phased array antennas operating in the X-band with a large frequency coverage to counter the effects of multi-path and jamming [105]. Maneuvering targets, or targets that are being engaged, will be tracked at a higher update rate than straight-moving targets or non-threatening targets. The multifunction capability has also lead to a new set of challenges. The jobs have to be scheduled smartly to harness the entire potential of a phased array radar. The scheduling requirements generally are specific to each radar system. Different approaches are based on neural networks,



decision theoretic, information theory and mathematical programming techniques, including linear, nonlinear and dynamic programming [107].

More recently, the concept of cognitive radar was introduced [108]. According to Haykin, there are three ingredients that are basic to the constitution of a cognitive radar: (i) intelligent signal processing, which builds on learning through interactions of the radar with the surrounding environment; (ii) feedback from the receiver to the transmitter, which is a facilitator of intelligence; and (iii) preservation of the information content of radar returns, which is realized by a Bayesian approach to target detection through tracking. Learning function means the cognitive radar has the inherent ability to sense its environment in a continuous manner and also refers to the ability of phased-array antennas to electronically scan the environment in a fast way. The learning procedure of the cognitive radar mimics the way in which the echo-location bat learns from its environment. Cognitive radars can be used in multifunction radars and non-coherent radar networks.

Bell considered the problem of radar waveform design for optimal detection and target information extraction when the targets are modeled as extended radar targets [109]. The first problem, that of waveform design for the optimal detection of radar targets that exhibit resonance phenomena, involves the design of radar waveforms and receiver-filters that maximize the output signal-to-noise ratio at the receiver-filter output under constraints on transmitted waveform energy and duration. The second problem deals with the design of radar waveforms which maximize the mutual information between an ensemble of extended targets and the receiver-filter output.

Bandwidth is a limited and costly resource and is needed by multiple applications. Bandwidth sharing between radar and communications was discussed in [110], [111]. Jackson et al. investigated the challenges of improving radar spectral efficiency with a view to sharing the bandwidth with other users [110]. Cooperative sensing was used to improve the feasibility of coexistence of radar and communications in [111].

Joint multi-modal RF-EO multi-target tracking using adaptive waveform design and control was discussed in [112]. A particle filter was used to track targets while incorporating agility to the multi-modal tracking system to improve the overall tracking performance. Waveform design for the RF sensor was used to increase the probability of detection and decrease the measurement error covariance. Multi-target tracking requires the RF sensors to maintain a large enough surveillance area to include all targets, while the range resolution is required to be small enough to provide accurate range measurements. As we need to consider limitations due to processing power in the overall system, there should be tradeoffs between the FOV and range resolution in designing the RF waveform, and similarly, between the FOV and angle resolution in adapting the EO control. Furthermore, as the RF sensors are distributed, they obtain different range-rate measurements from different aspects of the targets.

Most of the cited work has focused on diversifying the functions performed by radar. While adaptive waveforms have been researched, the idea of waveforms with different bandwidths has been not been considered. The idea of bandwidth optimization for a radar target scene had not been addressed. While bandwidth sharing between applications has been explored, no attempt



has been made to optimize this process. Most of the scheduling work has focused on phased array systems rather than scanning or rotational antenna systems.

The motivation for our research is the need to build a radar which can adapt itself based on the target scenario. The goal is to have a radar which is capable of multiple fields of view and selects the appropriate one based on the requirement of the current scene. This radar, henceforth referred to as multimodal radar, should start off with a low bandwidth waveform and then move to higher bandwidths if the situation demands it. This radar should be able to operate in a staring mode for sensing a target scene and a scanning mode to perform surveillance of a wide field.

The multimodal radar will use only as much bandwidth as required by the target scene. The unused bandwidth can be used by other applications. This also opens up the question of how much bandwidth is optimal for a target scene. Is it possible to optimize the distribution of bandwidth among different applications? With continuously changing target scenarios, the bandwidth used up by the radar also changes accordingly leading to a constantly changing optimization problem.

We also need to consider the problem of scheduling a multimodal radar. This is slightly different from a conventional radar since the multimodal radar requires multiple passes. Also, we consider a scanning rotational radar since it was used by us to perform the surveillance of a field.

## **6.2. Concept of Multimodal Radar**

We know that the resolution of radar increases with the increase in bandwidth. However bandwidth is a costly resource and sometimes the return from higher resolution is not worth the additional cost.

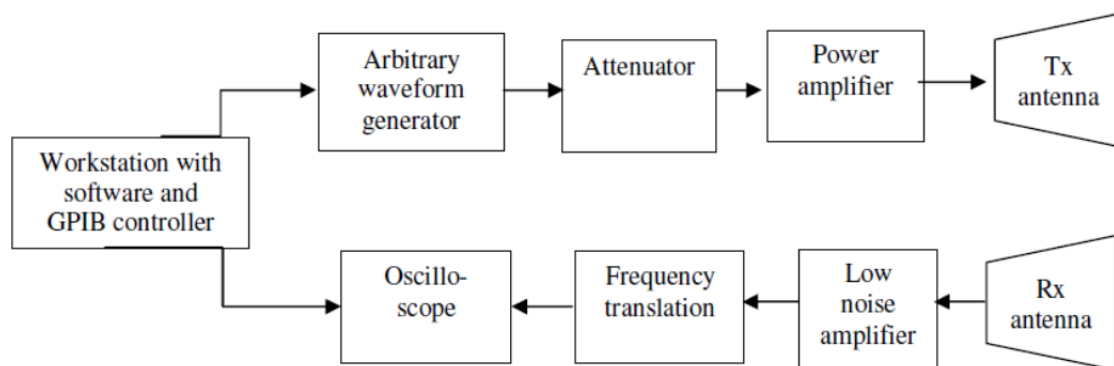
The target scenarios are constantly changing. Hence the optimum bandwidth required to extract information about the target scene also changes constantly. Hence, operating a radar at the maximum bandwidth at all times can be wasteful and it does not leave any bandwidth for other applications that may need it. Hence, we need a radar system which will only use as much bandwidth as is required for the target scene.

A multimodal radar along the above lines was first proposed by us. Initially, LRR profiles are formed using a low bandwidth waveform of 20 MHz. The detection threshold was arranged so that the potential targets would have a high probability of detection. HRR processing is then performed on selected range cells in which targets are declared. Thus the multimodal radar has the ability to provide target indication with a large range extent and can progressively switch to a narrow range extent mode for extracting recognizable target features. A hybrid waveform with a total bandwidth of 640 MHz is used for this purpose. This paper developed a simple test-bed using direct digital synthesizer to support multimodal requirements.

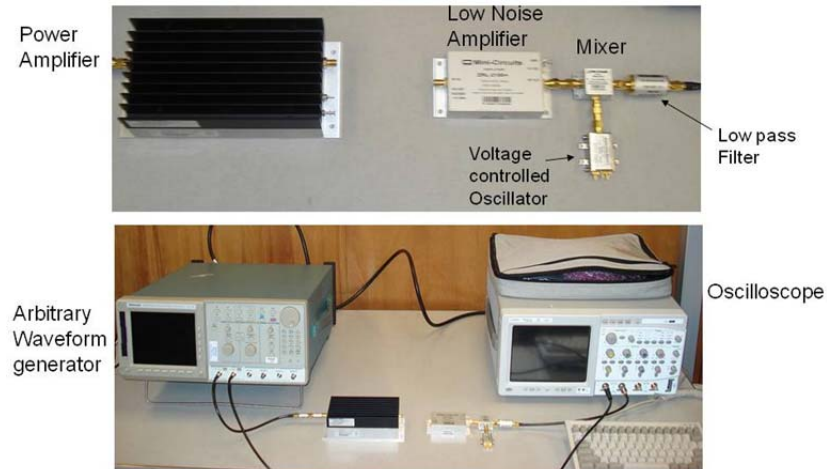
While Adler put forth the idea of dual resolution modes for radar [98], we carried the concept forward to multiple resolution modes where the radar switches from lowest resolution mode to the appropriate resolution mode based upon the target scene. References [113] and [114] are conference proceedings related to the multimodal radar.

### 6.2.1. Block Diagram and System Parameters

Figure 55 shows the block diagram of the multimodal radar and Figure 56 shows the components of the multimodal radar. The arbitrary waveform generator (AWG) enables generation of chirp waveforms of different bandwidths. The AWG operates at a maximum rate of 4 GSa/sec, making it possible to generate waveforms of frequencies up to 2 GHz. The test-bed radar also includes amplifiers, transmitting and receiving antennas, and a high-speed oscilloscope for recording the received signal. The equipment is shown in Figure 2. The power amplifier has a gain of around 30 dB and the maximum power output is +28 dBm (630 mW). The low noise amplifier has a typical gain of 25 dB and noise figure of 1.5 dB at 1 GHz. The mixer has a conversion loss of 6.5 dB at operating conditions. The frequency spectrum of the transmitted signal lies within the 1000–1640 MHz band. Frequency translation is performed to downconvert the received signal from the 1000–1640 MHz band to the 300–940 MHz range. The oscilloscope samples and records the return signal at 4 GSa/sec. The return is processed by software which decides whether further processing is required and what bandwidth waveform must be used for the next pass, if any. A workstation with a GPIB controller makes it possible for the software to control the AWG and the oscilloscope. Hence, the radar is capable of automatically making decisions about additional processing and required bandwidth. The radar system parameters are as shown in Table 7.



**Figure 55. Notional Block Diagram of the Multimodal Radar**



**Figure 56. Components and Equipment for the Multimodal Radar**

**Table 7. Radar System Parameters**

Parameter	Value
Radar waveform	Linear frequency modulated pulse
Radar bandwidth	40 – 640 MHz
Pulse width	16 $\mu$ s
Transmit power	Approx. 0.5 W
Maximum radar range	Approx. 25 m

### 6.2.2. LRR and HRR Profiles

Initially, the AWG transmits a low bandwidth waveform (40 MHz) and sweeps the range extent searching for targets. Range resolution  $\Delta R$  is given by  $c/2B$  where  $c$  is the speed of light and  $B$  is the bandwidth. Thus a bandwidth of 40 MHz corresponds to a range resolution of 3.75 m. The return is compared to an adaptive range-dependent detection threshold which takes into account the fact that the response from targets will weaken as the range increases. This gives us the LRR gates with a high probability of the existence of potential targets. The multimodal radar now restricts its attention to these LRR gates where the threshold is exceeded. HRR imaging begins on these identified LRR gates with the 80-MHz bandwidth waveform (1.875-m resolution). Imaging stops if the desired range resolution is obtained on a particular LRR gate to identify existing targets, else it continues with the next higher bandwidth (160 MHz  $\rightarrow$  320 MHz  $\rightarrow$  640 MHz). Range resolution is progressively enhanced until a minimum separation (in dB) is met between the peaks and its neighboring cells. This minimum separation may be decided based upon the required resolution and the expendable bandwidth. A 3-dB separation was used for most of the experiments performed by us. A higher separation would obviously lead to more passes and higher bandwidth usage by the multimodal radar. Thus, the multimodal radar

continues to look at potential targets with narrower range extents until the desired resolution is obtained to detect target presence. Table 8 summarizes the various bandwidths used by the multimodal radar and their corresponding range resolutions.

**Table 8. Bandwidth and Resolution for Each Pass of the Multimodal Radar**

Pass	Bandwidth (MHz)	Resolution (m)
1	40	3.75
2	80	1.87
3	160	0.93
4	320	0.46
5	640	0.23

### 6.2.3. Field Setup

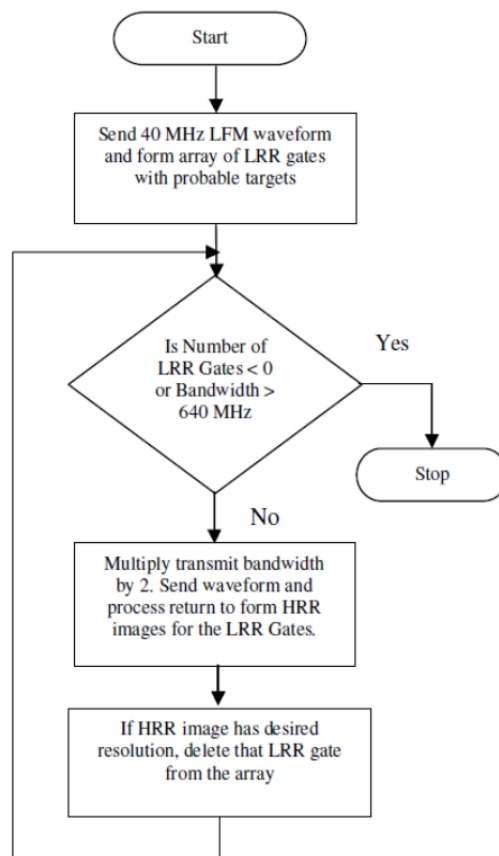
The measurement set up is shown in Figure 57. Trihedral corner reflectors with square faces of length 0.6 m were used as targets whose radar cross section (RCS) was computed as  $57.8 \text{ m}^2$  (+17.6 dBsm). We consider a range of 37.5 m which is slightly greater than the maximum radar range. This results in 10 LRR gates each of extent of 3.75 m. The RCS values of the targets are normalized with respect to the target with the highest RCS. The algorithm continues with higher resolution passes until 3-dB separations are obtained between the peaks and their neighboring cells.



**Figure 57. Field Measurement Set Up for Multimodal Radar**

#### 6.2.4. Operation in Staring Mode

Figure 58 shows the flowchart for operation of the multimodal radar. It is assumed that the frequency band of 1–1.64 GHz is available for use. A list of LRR gates is formed using a LFM waveform of bandwidth 40 MHz. Then the bandwidth is increased in powers of two until the desired resolution is obtained for each HRR profile. The simulation may require different number of passes for different LRR gates. When the desired resolution is obtained for a LRR gate, it is dropped from the list and the sensing continues with the remaining LRR gates. The multimodal radar makes efficient use of the spectrum, utilizing lower bandwidths initially, and using the higher bandwidths only if required. This leaves the unused bandwidth for use in other applications.

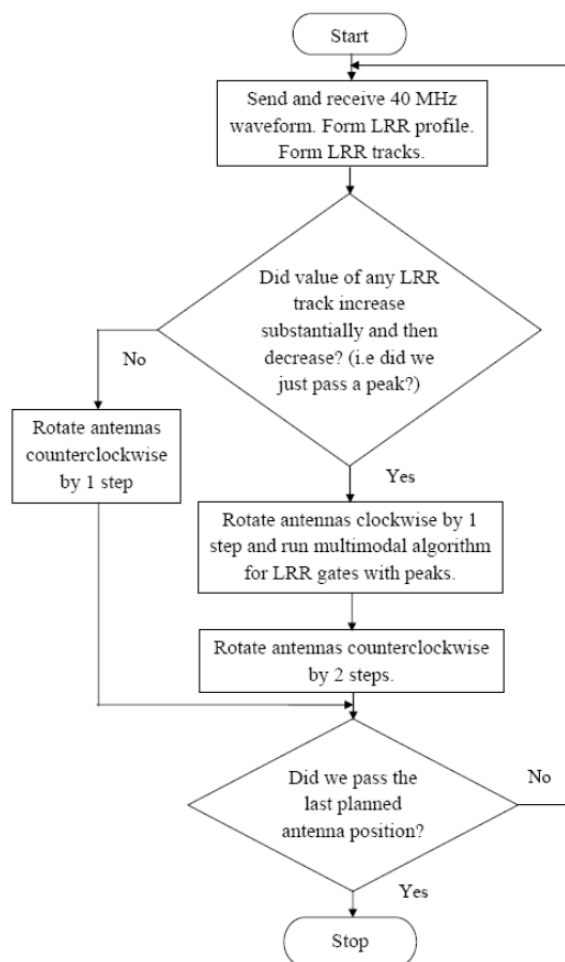


**Figure 58. Flowchart for Operation of Multimodal Radar.**

#### 6.2.5. Operation in Scanning Mode

The multimodal radar was also programmed to scan a field of targets. We start detection with a low bandwidth waveform. The antennas are mounted on a plate which can rotate in the azimuthal

plane. The idea is to rotate the antenna in small steps of around 15–20°. When a target is detected, the radar starts increasing the bandwidth in steps of two until the desired resolution is obtained. The first LRR profile is taken at  $\phi = 0$  and subsequent profiles are taken after rotating the antennas counter-clockwise each time by a fixed angle, henceforth referred to as one step. The LRR profiles are stored to form LRR tracks for each LRR gate. The number of LRR gates would be finite, based on the range of the radar. In the LRR tracks, we look for peaks since they would signify the presence of a target. Also we would like to use higher bandwidth waveforms from the position which is closest to the line joining the center of antennas and the target. For these reasons, we keep rotating until a peak is passed and then the antennas are rotated clockwise (in the opposite direction) by one step to run the multimodal algorithm on the target scene. Only the tracks which showed peaks would be looked at by the higher bandwidth waveforms. Once the desired resolution is obtained, the rotate-and-scan-process continues. Figure 59 shows the flowchart for operation of the radar when scanning in this fashion. The entire setup is automated so that the scanning process can operate without any user intervention. The antenna angle measurements are read using a USB data acquisition device.



**Figure 59. Operation of Multimodal Radar when Scanning a Field**

### 6.3. Results from Theory and Simulations

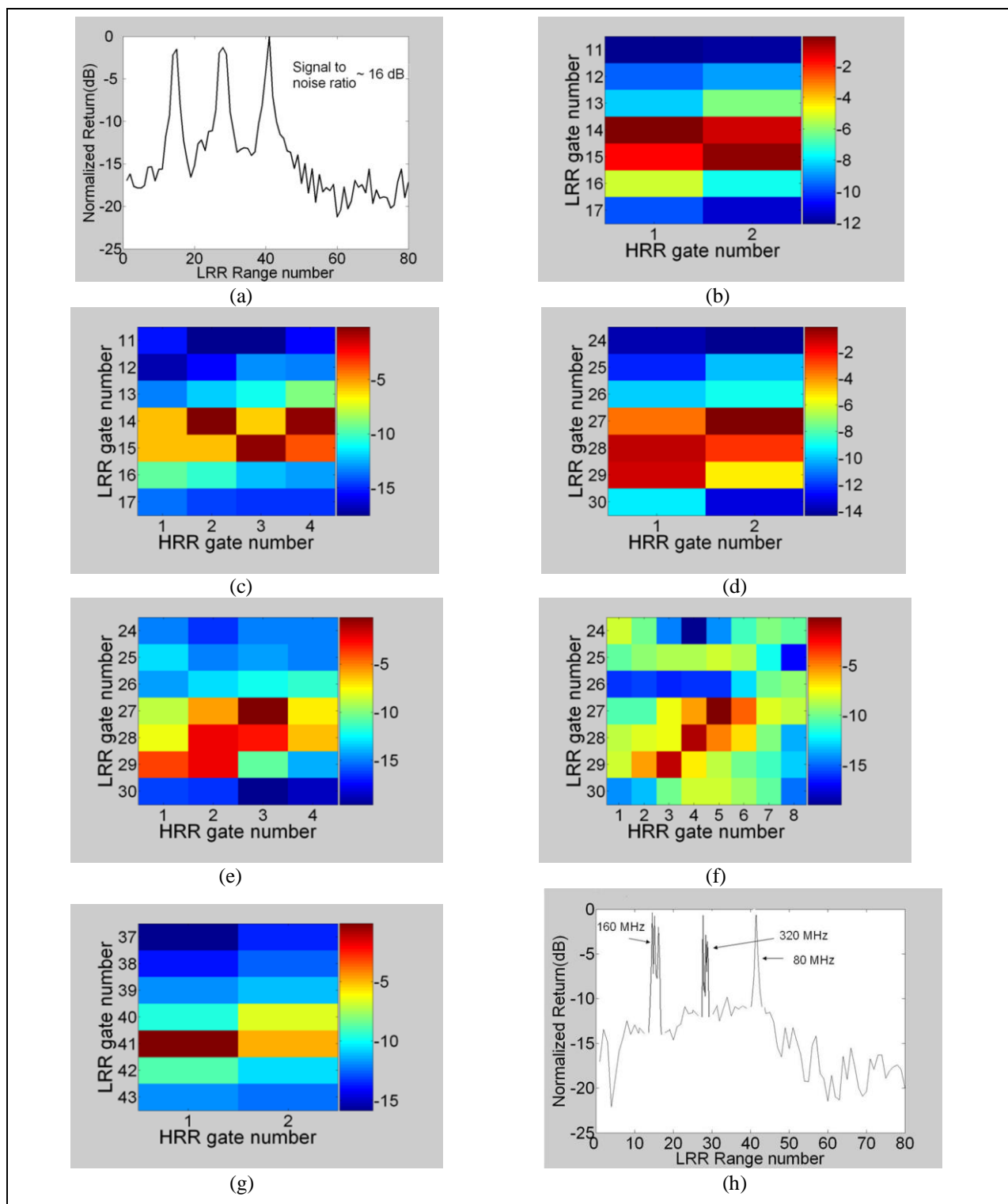
#### 6.3.1. Initial Simulation Results

The following show the results obtained using simulations in MATLAB<sup>®</sup>. These were tried before the field experiments were conducted and acted as a preliminary proof of concept of the multimodal radar. We considered a maximum range of 300 m, which results in 80 LRR gates each of 3.75m extent. For this simulation, there exist targets as shown in the Table 9. The radar cross sections (RCSs) of the targets are normalized with respect to the target with the highest RCS. Targets of higher RCS are used at higher range to give sufficient response to be detectable by the system.

**Table 9. Target Scenario 1**

Target Number	Range (m)	LRR Gate	Relative RCS
1	50	14	0.01
2	52	14	0.01
3	55	15	0.01
4	100	27	0.15
5	103	28	0.15
6	106	29	0.15
7	151	41	1

These targets are centered around LRR gates 14, 27 and 40. These three LRR gates are observed in the LRR profile as seen in Figure 60(a). The LRR resolution is not low enough to distinguish between the individual targets with a set. Hence the simulation proceeds to Pass 2 with 80 MHz while restricting its focus to these LRR gates. HRR images are generated for each of the selected LRR gates. For the HRR profiles, the y-axis is in increments of LRR index, and the x-axis is in increments of HRR index. For the bandwidth of 80 MHz, each LRR gate has 2 HRR cells. As the bandwidth increases, the number of HRR cells per LRR gate increases. The HRR images are normalized and the cell values are in dB. Simulation continues for each of the LRR gates until specified amount of separation is obtained between the peaks and their neighboring cells. For this particular example, 3 peaks were taken into consideration for each HRR profile. The peaks are discarded if they are not significantly higher than the mean of all the cells. During Pass 2, 3-dB separations between the peaks and neighboring cells are achieved for LRR gate 41 as seen in Figure 60(g) and hence the simulation for this gate stops. Meanwhile more resolution is still desired for LRR gates 14 and 27 and the simulation for these gates proceeds to Pass 3 with 160 MHz. The desired resolution for LRR gate 14 is achieved in pass 3 as seen in Figure 60(c) and this simulation stops. Meanwhile the simulation for LRR gate 27 continues to Pass 4 with 320 MHz and finally stops there. Thus the multimodal radar is able to look at different target scenes with different resolutions.



**Figure 60. Imaging Results for a Simulated Target Scenario 1**

**(a) Pass 1. (b) Pass 2, LRR Gate 14. (c) Pass 3, LRR Gate 14. (d) Pass 2, LRR Gate 27. (e) Pass 3, LRR Gate 27. (f) Pass 4, LRR Gate 27. (g) Pass 2, LRR Gate 41. (h) Resulting resolution for each LRR Gate**



## 6.3.2. Bandwidth Optimization

### 6.3.2.1. Optimization Problem

The drawback of the aforementioned method is that it requires multiple passes. Each pass consumes certain time and resources and if a target scene is intricate enough to require maximum bandwidth, then the multimodal radar utilizes more resources than traditional radar operating at that bandwidth. What is desirable is to look at a target scene and arrive at the optimum bandwidth required for that target scene and directly use it for pass two of the multimodal radar. Thus the cost function for our optimization problem is the bandwidth which is same as the optimization variable. The separation between the peak and its neighboring cells can serve as the constraints. We would like to have 3-dB separations between the target peak and its neighboring cells. This should be adequate to provide a good representation of the range-related variability within the target scene. Using more separation may increase the required bandwidth significantly. Using lesser separation may not give the required resolution to discern the target scene.

The optimization problem for the bandwidth of the multimodal radar can be stated as

$$\begin{aligned}
 &\text{minimize} && x \\
 &\text{subject to} && f_1(x) \geq 3 \text{ dB} \\
 & && f_2(x) \geq 3 \text{ dB} \\
 &\text{where} && x = \text{bandwidth used for a target scene} \\
 & && f_1 = \text{separation between peak and neighboring cell towards the radar} \\
 & && f_2 = \text{separation between peak and neighboring cell away from the radar}
 \end{aligned}$$

The above constraints can be rewritten as  $3 - f_1 \leq 0$ ,  $3 - f_2 \leq 0$ .

The above constraints will generally not be convex. As the bandwidth changes, the size and position of the range gates change and the values of  $f_1$  and  $f_2$  change in a disorderly fashion. However, if the gate containing the peak target is forced to be centered on it, then the above constraints can be approximated by convex functions. The accuracy with which the functions can represent the above constraints will greatly influence the correctness of the results.

The Lagrangian duality method was used to solve the problem [115]. The basic idea is to take the constraints into account by augmenting the objective function with a weighted sum of the constraint functions.

$$L(x, \lambda) = x + \sum_{i=1}^2 \lambda_i (3 - f_i) \quad (228)$$

where  $\lambda_i$  is the Lagrange multiplier associated with the  $i^{\text{th}}$  inequality constraint. The Lagrange dual function  $g$  is the minimum value of the Lagrangian over  $x$ . If  $D$  represents the domain of the problem, then the dual function [115] is expressed as

$$g(\lambda) = \inf_{x \in D} L(x, \lambda) \quad (229)$$

The maximum of the dual function  $g$  gives a lower bound on the optimum value of  $x$ . For strong duality (which holds for convex problems), the maximum of the dual function is exactly equal to the optimum value of  $x$ . The following shows an example where the constraints are approximated by convex quadratic functions.

$$3 - f_1 \cong a_1 x^2 + b_1 x + c_1 \quad (230)$$

$$3 - f_2 \cong a_2 x^2 + b_2 x + c_2 \quad (231)$$

The constants in the above equations depend on the target scenario. The values for  $f_1$  and  $f_2$  for certain values of bandwidths are found using simulations and then appropriate quadratic functions are used to represent the constraints. Putting (229) and (230) in (228), the Lagrangian becomes

$$L(x, \lambda) = x + \lambda_1 (a_1 x^2 + b_1 x + c_1) + \lambda_2 (a_2 x^2 + b_2 x + c_2) \quad (232)$$

To minimize  $L$  over  $x$ , set gradient equal to zero.

$$\begin{aligned} \frac{\partial L}{\partial x} &= 1 + \lambda_1 (2a_1 x + b_1) + \lambda_2 (2a_2 x + b_2) = 0 \\ \Rightarrow x &= \frac{-(1 + b_1 \lambda_1 + b_2 \lambda_2)}{2(a_1 \lambda_1 + a_2 \lambda_2)} \end{aligned} \quad (233)$$

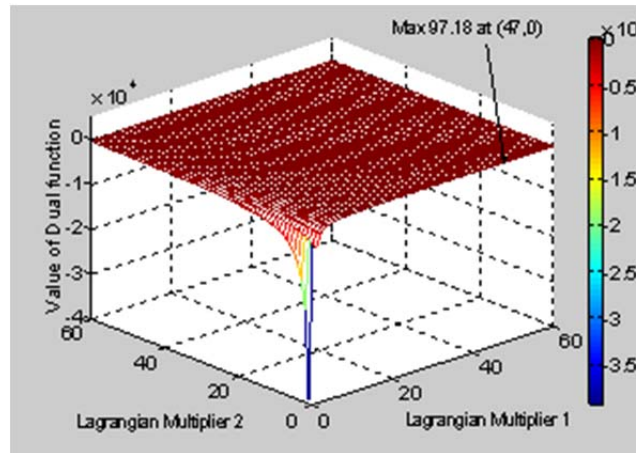
Plugging (232) in (231) gives the dual function  $g(\lambda)$ . For a convex problem, the maximum of the dual function  $g$  gives the optimum value of  $x$ .

### 6.3.2.2. Examples

Target scenario 2 is shown in Table 10. There are two targets separated by 2.5 m. A range resolution value of 2.5 m or better will be able to resolve the two targets, but this is only assuming that the point spread function is an ideal impulse. In reality, the point spread function is usually a Gaussian or a sinc-function having energy beyond its 3-dB point, which may cause a larger target to obscure a closer but smaller target. Thus, to be on the safe side, a resolution value of half of the minimum target separation is used to ensure capturing targets of all sizes and make target identification possible. Thus, in this case, a resolution of about 1.25 m would be ideal. Simulations were performed to find sample values for the constraints and they were represented by convex quadratic functions. The dual function was obtained and its values for different values of Lagrangian multipliers are shown plotted in Figure 61. The maximum of the dual function has a value of 97.18 MHz, which corresponds to a range resolution of 1.54 m. This is close to the desired 1.25-m resolution, indicating that our technique works well.

**Table 10. Target Scenario 2**

Target Number	Range (m)	Relative RCS
1	8.2	1
2	10.7	1

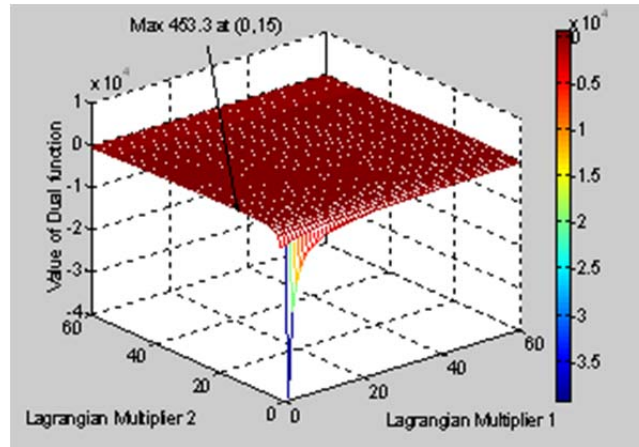


**Figure 61. Dual Function for Target Scenario 2.**

Another target scenario 3 is shown in Table 11. The minimum separation is 0.6 m, calling for a resolution of about 0.3 m. The maximum of the dual function has a value of 453.3 MHz corresponding to a range resolution of 0.33 m, which is again close to the desired value. The dual function is shown in Figure 62. The multimodal radar required a bandwidth of 640 MHz for this scenario. This is justified since the optimum bandwidth is higher than the Pass 4 bandwidth of 320 MHz.

**Table 11. Target Scenario 3**

Target Number	Range (m)	Relative RCS (sq. m)
1	11.5	1
2	12.1	1

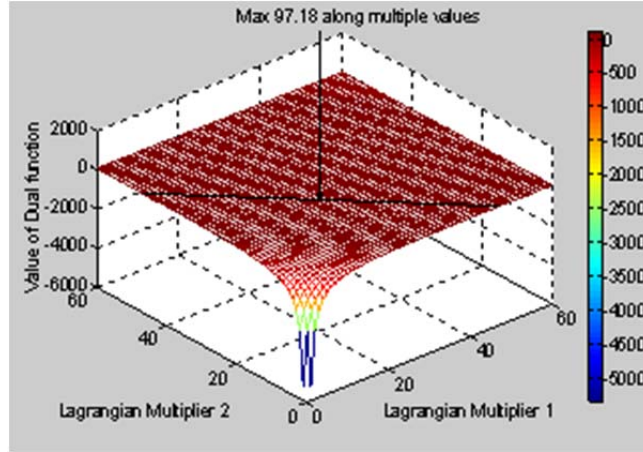


**Figure 62. Dual Function for Target Scenario 3.**

In the above two examples, we see that the maximum occurs at a value where either  $\lambda_1$  or  $\lambda_2$  is zero. This signifies that one of the constraints is dominant and is masking the other constraint. Hence a symmetrical target scenario was tried as shown in Table 12. This is similar to target scenario 3 except that another target is added to make the scene symmetrical. The optimum bandwidth is 97.18 MHz, which is the same as was obtained for target scenario 3. Here the maximum occurs at multiple values of pairs of  $(\lambda_1, \lambda_2)$ , some of which have both  $\lambda_1$  and  $\lambda_2$  non-zero. The dual function for this target scenario is shown in Figure 63.

**Table 12. Target Scenario 4**

Target Number	Range (m)	Relative RCS
1	8.2	1
2	10.7	1
3	13.2	1

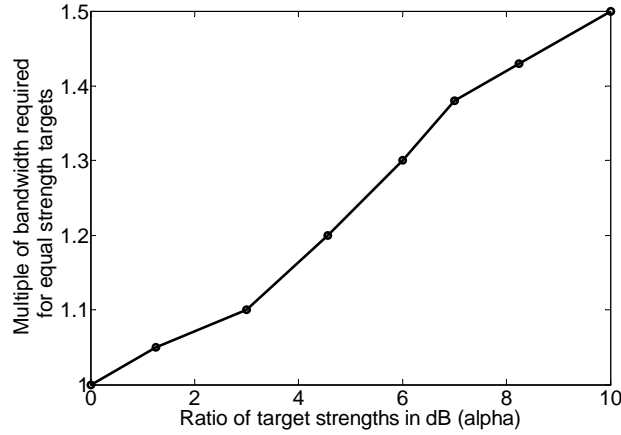


**Figure 63. Dual Function for Target Scenario 4**

### 6.3.3. Bandwidth Requirement for Targets Of Unequal Strength

For the previous discussion, we assumed that the targets were of equal strength. However, if the targets vary greatly in strength, a resolution value of half of the target separation is no longer accurate. In [116], the common definition used for measuring range resolution for equal strength targets was modified for targets of unequal strength. Simulations were carried out to find a relationship between the strength difference and the bandwidth requirement. It is shown in Figure 64. The ratio of target strengths is expressed in dB. The change in bandwidth is expressed as the multiple of bandwidth required for equal strength targets. The relationship is almost linear. Hence for targets of unequal strength, the bandwidth optimization problem can be stated as

$$\begin{aligned}
 &\text{minimize} && x \\
 &\text{subject to} && f_1(x, \alpha) \geq 3 \text{ dB} \\
 &\text{where} && x \text{ is the bandwidth used for a target scene,} \\
 & && \alpha \text{ is the ratio of target strengths in dB } (\alpha \geq 0) \\
 & && f_1 \text{ is the separation in dB between the cell containing the weaker target and the} \\
 & && \text{neighboring cell towards the stronger target.}
 \end{aligned}$$



**Figure 64. Dependence Of Bandwidth Required for Resolution on Target Strength Ratio**

To solve the above problem, we first solve the problem using  $\alpha = 0$  and arrive at the solution  $\beta$  using the procedure introduced before. The actual bandwidth  $\beta_{\text{req}}$  for the problem can then be arrived at using the following expression

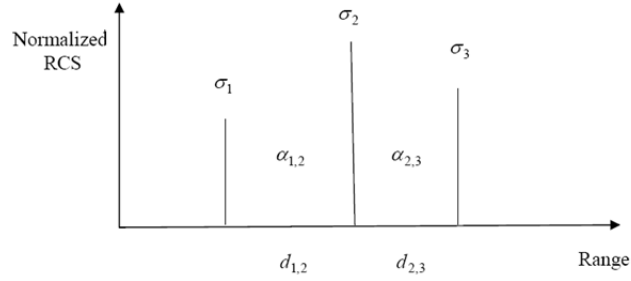
$$\beta_{\text{req}} = \beta \left( 1 + \frac{0.5\alpha}{10} \right) = \beta \left( 1 + \frac{\alpha}{20} \right) \quad (234)$$

Let us consider an arbitrary target shown in Figure 65 and calculate the bandwidth required for its detection. For each pair of adjacent scattering centers, we take the ratio of the RCS of the stronger one to the weaker one to get  $\alpha$  in dB. We separately look at the bandwidth required for resolving each pair of adjacent scattering centers. Let us first look at the pair  $(\sigma_1, \sigma_2)$ . We approximate the expected value of bandwidth based on the theory of point spread function. If  $\sigma_1$  and  $\sigma_2$  were equal, the resolution required would have been approximately  $d_{1,2}/2$ . This corresponds to a bandwidth of  $\beta_{1,2} = c/d_{1,2}$ . When  $\sigma_1$  and  $\sigma_2$  are unequal, this can be adjusted to be equal to

$$\beta_{1,2 \text{ req}} = \frac{c}{d_{1,2}} \left( 1 + \frac{\alpha_{1,2}}{20} \right) \quad (235)$$

Looking at the other pairs of scattering centers, for a total of  $n$  scattering centers, the bandwidth required can be expressed as

$$\beta_{\text{req}} = \max_{i \in (1, n-1)} \left[ \frac{c}{d_{i,i+1}} \left( 1 + \frac{\alpha_{i,i+1}}{20} \right) \right] \quad (236)$$



**Figure 65. Scattering Centers of an Extended Target.**

When we consider just the pair of adjacent scattering centers, we neglect the effect of other scatterers. This can be justified from the fact that if a scatterer was close enough to make a substantial contribution, then the bandwidth required to resolve that scatterer from its neighbor would be much higher in the expression above.

#### **6.3.4. Observations from Simulations**

In this section, we look at certain results obtained from simulations. We also try to understand how the number of passes required by the multimodal radar would change with respect to external conditions. These results are based on extensive number of repetitions of the multimodal algorithm while we change some parameter such as distance between targets and SNR.

##### **6.3.4.1. Receiver Operating Characteristics**

Receiver operating characteristic (ROC) of any radar is critical for its analysis. Simulations were performed to generate the ROC curves for the multimodal radar. Before we delve into the results, let us see how we may define the probability of detection ( $P_d$ ) and probability of false alarm for a multimodal radar ( $P_f$ ).

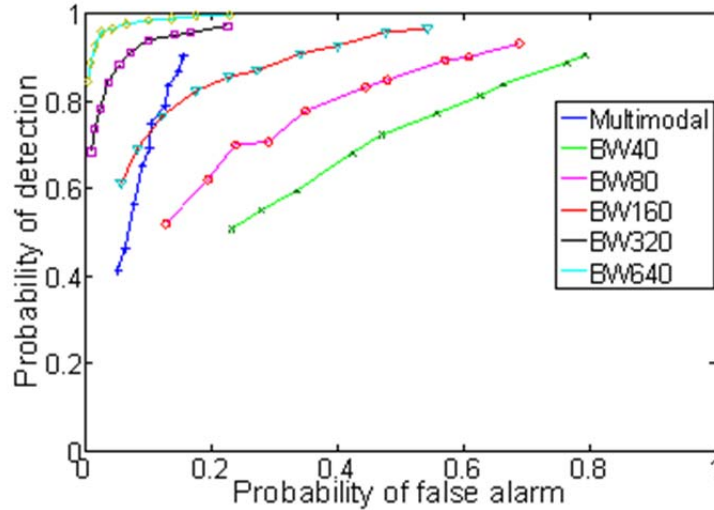
$$P_d = \frac{\text{Number of targets correctly forecast in every pass}}{\text{Total number of targets}} \quad (237)$$

$$P_f = \frac{\text{Number of targets that were falsely forecast in all passes}}{\text{Total number of non-target cells}} \quad (238)$$

We generate multiple targets whose ranges are generated randomly. The RCS of targets is assumed to be proportional to the fourth degree of the range. The multimodal algorithm is run and different LRR gates will end the simulation after different number of passes. It may be noted that  $P_f$  varies depending on the number of passes executed. It is proposed that the  $P_f$  for the multimodal radar be taken as the average of all of these.

$$P_f = \frac{P_{f(2)} + P_{f(3)} + P_{f(4)} + P_{f(5)}}{4} \quad (239)$$

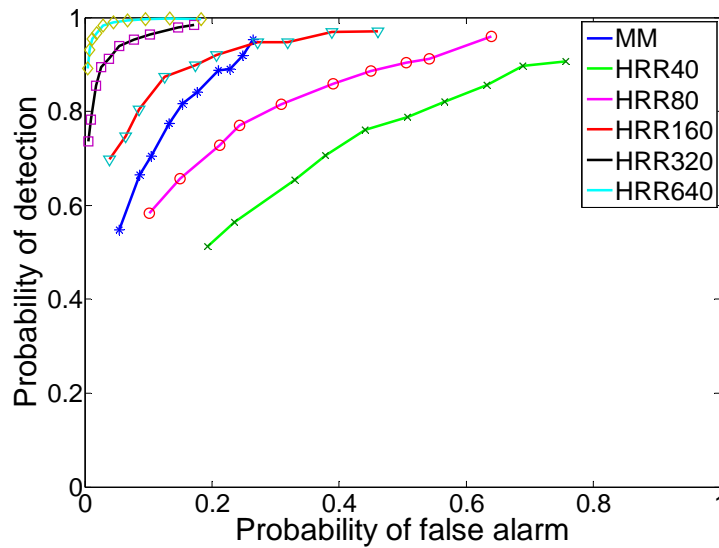
Figure 66 shows the ROC of the multimodal radar and compares it with the ROC plots of single pass radars of various bandwidths. The signal-to-noise ratio (SNR) in the simulations is around 5–7 dB. As expected, for a given  $P_f$ , the  $P_d$  improves as bandwidth increases. We also note that the  $P_f$  of the multimodal radar does not deteriorate too rapidly with the increase in the  $P_d$ , as is seen for the single pass radars. This is expected since the existence of multiple passes reduced the chances of false alarm as compared to single pass radars. Also, the  $P_d$  of the multimodal radar is limited by the  $P_d$  of its LRR pass. This follows from the fact that the LRR pass is the first pass in the multimodal radar and any target missed in this pass goes undetected in the higher bandwidth passes which focus only on the LRR gates found earlier. Hence, the  $P_d$  of the multimodal radar may be improved by selecting an initial LRR pass with higher bandwidth.



**Figure 66. ROC of Multimodal Radar.**

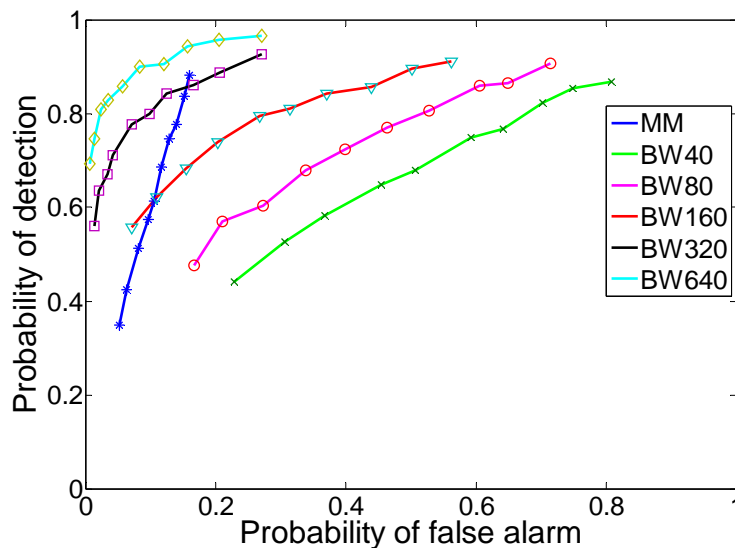
The simulation was repeated with the multimodal radar operated with a LRR pass of 80 MHz instead of the usual 40 MHz bandwidth. This is shown in Figure 67. Here the  $P_d$  values of the multimodal radar are better than the single pass radar of bandwidth 40 MHz (BW40) but lower than the single pass radar of bandwidth 80 MHz (BW80). This improvement in  $P_d$  is however accompanied by worse values of  $P_f$ , especially at lower detection thresholds.





**Figure 67. ROC When Multimodal Radar has LRR Pass with 80 MHz Bandwidth.**

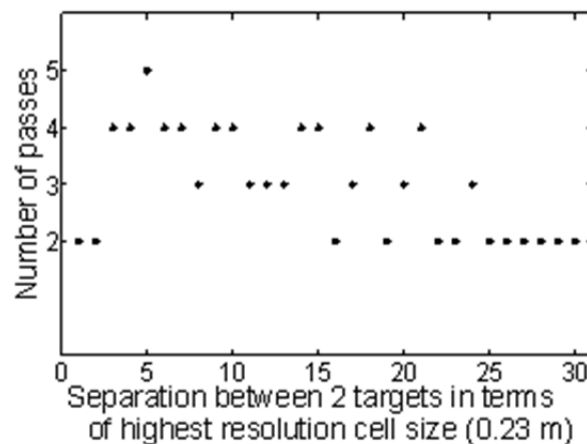
The above simulations were carried out with the RCS values for targets assigned according to their range so that all targets would give an approximately uniform return. This condition was relaxed to allow for a 10 dB variation in the strength of the targets. The LRR pass had a bandwidth of 40 MHz. The result is shown in Figure 68.



**Figure 68. ROC When Targets are Varying In Strength**

#### 6.3.4.2. Effect of Separation Between Targets

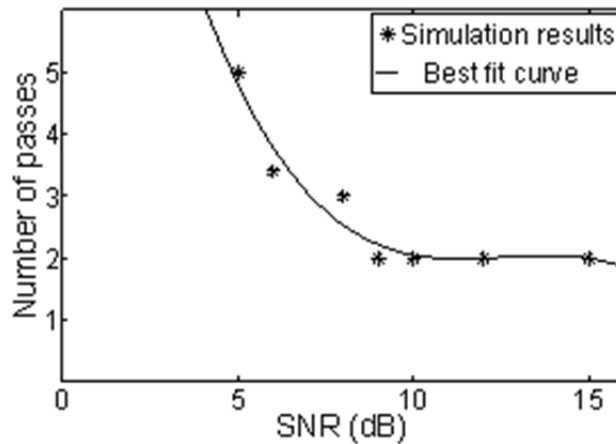
We explore the effect of separation between the targets. We consider two targets and note the number of passes required as the separation between them is gradually reduced. The result is plotted in Figure 69 where the X-axis is in terms of the highest resolution cell size (0.23 m). The number of passes required generally increases as the separation between the targets is reduced. Hence, more bandwidth would be utilized as the distance between two targets is reduced. This suggests that as the target scene becomes more intricate, the multimodal radar would require a greater number of passes. The number of significant peaks observed in Pass 2 HRR images can also provide an indication of an intricate target scene. An important consideration is whether an attempt should be made to skip some intermediate bandwidths and directly advance to a higher bandwidth for such target scenes. Figure 69 shows that when the target separation is too low, the multimodal radar took only 2 passes. This happened in Pass 2 because the two targets appeared as a single target.



**Figure 69. Variation In Required Number of Passes as a Function of Target Separation**

#### 6.3.4.3. Effect of SNR

We also simulated the effect of SNR on the performance of the multimodal radar. The target scenario was kept the same while the SNR was varied. It was observed that a higher number of passes is required to resolve a target scene as the SNR is decreased. The result is plotted in Figure 70. This follows from the fact that better separations are obtained between the peaks and their neighboring cells at higher SNRs. As the SNR reduces, more passes are required to get adequate separation. At high SNRs, the number of passes levels off to a value of two for the scenarios under consideration, and starts to dip at an SNR of 15 dB. It is possible that for certain scenarios at high SNR values, a single pass may be adequate to resolve the target scene.



**Figure 70. Variation in Required Number of Passes as a Function of SNR**

### 6.3.5. Sharing Bandwidth Between Radar and Communications

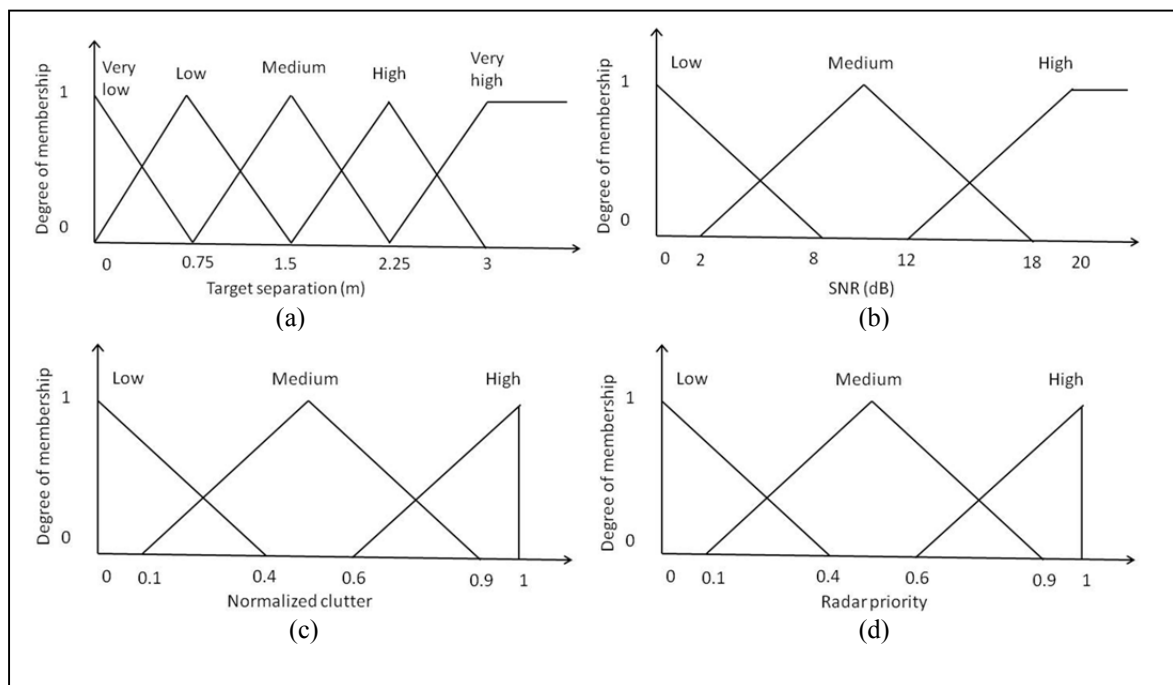
Sometimes we have limited bandwidth which needs to be shared among multiple applications. Since the multimodal radar can work with different bandwidths, it is well suited to these scenarios. We consider the problem of optimally share the bandwidth between radar and communications. We shall use multi-objective optimization techniques for this purpose.

#### 6.3.5.1. Determining Radar Priority Using Fuzzy Logic

We need a method to arrive at radar priority before we can apply optimization techniques. We use a method similar to [117] for this purpose. The mentioned paper develops an adaptive prioritization assignment, fuzzy-reasoning-based algorithm for ranking sectors of surveillance in dynamically changing tactical environments. The priority of surveillance sectors was assessed using aspects such as rate of detection of new targets and number of threatening targets. For our case, the linguistic variables used to determine radar priority are target separation, signal-to-noise (SNR) ratio and clutter. The membership functions for these variables are shown in Figure 71. The linguistic variable target separation is the fuzzified value of the distance between point targets. The distances correspond to the range resolutions of the various bandwidth waveforms used by the multimodal radar seen before. Since target separation is the single most important factor, this has been designed with more fuzzy values compared to other linguistic variables. Clutter is expressed on a normalized scale of 0 to 1. The membership functions for SNR and radar priority are pretty straightforward. The fuzzy rules provide a mapping from the input linguistic variables to the output of radar priority. We see that target separation features in all the rules used by the system. The centroid method is used for defuzzifying the output.

The following rules are used by the fuzzy inference system:

- If target separation is very low, radar priority is high.
- If target separation is very high, then radar priority is low.
- If target separation is medium and SNR is low, then radar priority is high.
- If target separation is medium and clutter is high, then radar priority is high.

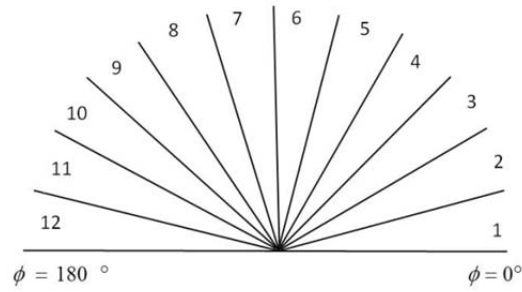


**Figure 71. Membership Functions**

**(a) Target separation (b) SNR (c) Normalized clutter (d) Radar priority**

### 6.3.5.2. Target Scenario 5

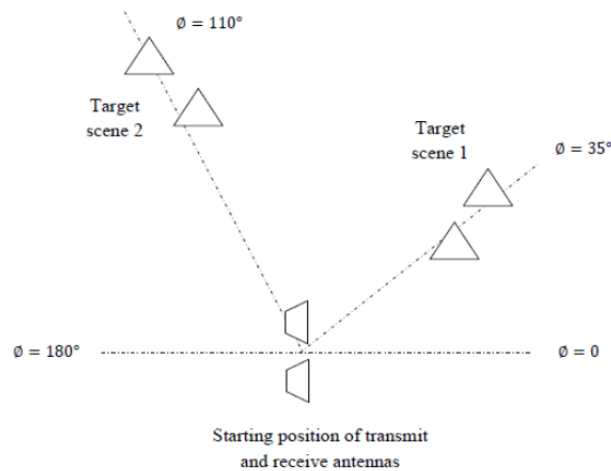
We consider the scene in front of the radar to be divided into  $15^\circ$  wide sectors as shown in Figure 72. The surveillance target scenario 5 being considered for bandwidth sharing and optimization is shown in Table 13. It is also shown diagrammatically in Figure 73. We start scanning from Sector 1 and move towards Sector 12. The communications scenario is assumed to be constantly changing and as shown in Table 14 corresponding to the instant when each sector is being scanned.



**Figure 72. Target Space Divided into Sectors**

**Table 13. Target Scenario 5**

Target Scene Number	(degrees)	Sector	Range (m)	Normalized Clutter	SNR (dB)
1	30	3	8.2	0.7	8
	30	3	10.7		
2	110	8	11.5	0.5	14
	110	8	12.1		



**Figure 73. Target Scenario 5**

Using MATLAB fuzzy logic toolbox, we arrive at priorities for radar for surveillance of the various sectors. The SNR and normalized clutter for non-target sectors are assumed to be 10 dB and 0.5 respectively. The results are shown in Table 15. Now that we have priorities for both radar and communications, we can proceed with the multi-objective optimization process.

**Table 14. Communications Priority**

Sector	Communications scenario
3	Large amount of data needs to be transmitted (Priority 0.8)
8	Medium amount of data needs to be transmitted (Priority 0.5)
Others	Medium amount of data needs to be transmitted (Priority 0.4)

**Table 15. Computed Radar Priority**

Sector	Radar priority
3	0.17
8	0.82
Elsewhere	0.13

### 6.3.5.3. Multi-Objective Optimization

Using the calculated values of radar priority and the assumed values for communications priority, we perform multi-objective optimization. The total available bandwidth is taken to be 640 MHz. This has to be shared optimally between radar and communication tasks.

The multi-objective problem can be defined as:

Maximize  $F(\bar{x}) = [F_1(\bar{x}) \ F_2(\bar{x})]$

subject to  $f_1(x_1) \geq 3 \text{ dB}$

$f_2(x_1) \geq 3 \text{ dB}$

$x_1 + x_2 \leq 640 \text{ MHz}$

where  $\bar{x} = [x_1 \ x_2]$

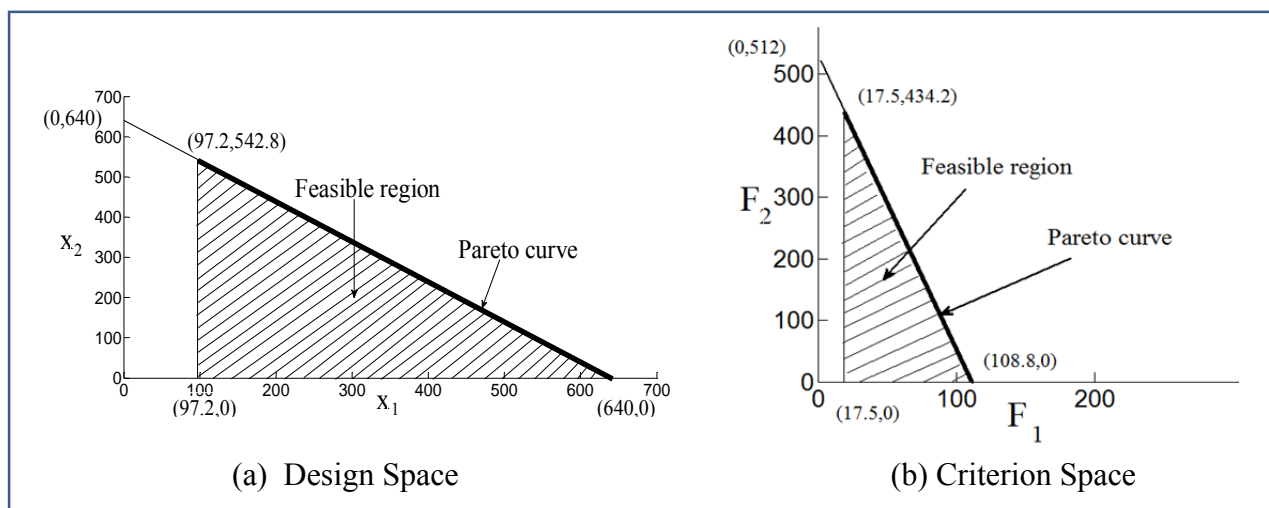
$x_1 = \text{Bandwidth used for radar}$

$x_2 = \text{Bandwidth used for communications}$

The 3-dB constraints ensure adequate resolution between the radar peak and its neighboring cells. The parameters of the problem vary for each sector. Using the aforementioned priorities as weights, we arrive at the following objective functions for Sector 3:

$$\begin{aligned} F_1(\bar{x}) &= 0.17x_1 \\ F_2(\bar{x}) &= 0.82x_2 \end{aligned} \tag{240}$$

For these parameters, the design space and criterion space are shown in Figure 74. The radar resolution constraints effectively place a lower bound on the radar bandwidth. This was calculated using the multimodal radar optimization process discussed before.

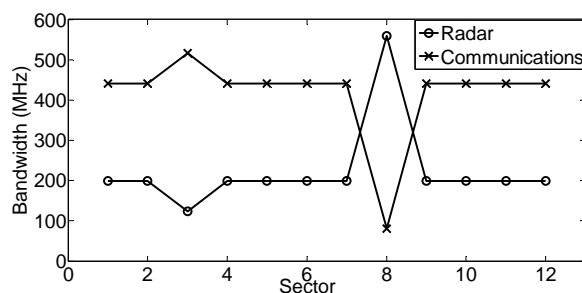


**Figure 74. Multi-Objective Optimization Solution for Sector 3**

Minimizing the Euclidean distance to the utopia point [118], the solution comes out to be:

$$\begin{aligned} x_1 &= 123.5 \text{ MHz} \\ x_2 &= 516.5 \text{ MHz} \end{aligned} \quad (241)$$

Solving for each of the sectors gives us the solution shown in Figure 75. We see that the bandwidth is appropriately allotted based on the requirement of radar target scene and communications. Thus the radar is assigned high bandwidth to allow higher resolution for Sector 8. Communications is assigned high bandwidth in Sector 3 as per the high priority.



**Figure 75. Multi-Objective Optimization Solution for the Problem Scenario.**

### 6.3.6. Radar Scheduling

Radar scheduling is important to make the best use of available resources. We investigate the scheduling of the multimodal radar for surveillance and range tracking. The antennas are rotated

using TTL/relay-controlled DC motor. For our analysis, we neglect the radar dwell time, since it is negligible compared to the return processing time and time required to rotate the antenna from one sector to another. The parameters used for the following simulations are shown in Table 16.

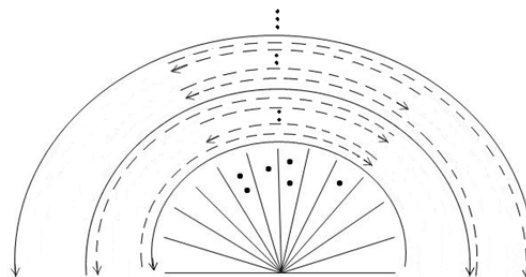
**Table 16. Simulation Parameters**

Parameter	Value (sec)
Radar return processing time	2
Time to rotate antenna to adjacent sector	1

We maintain a minimum number of visits for targets being tracked while allowing for non-uniform samples. Tracking using non-uniform sampling has been investigated before in [119], [120]. We assume slow moving targets which continue to remain in the same sectors for the duration of the simulation. We differentiate between scenarios having targets in the center and targets on the edges.

#### 6.3.6.1. Targets in the Center

For such target scenarios, the algorithm shown in Figure 76 is used. Partial passes are used close to the center to increase the number of visits to sectors with targets. Only the sectors with targets are sampled in passes shown with dotted arcs. All sectors are sampled in passes shown with full arcs. After one complete cycle, the number of visits is measured for the sectors with targets and the number of partial passes between full passes is increased if required. Surveillance is performed for the sectors with no initial targets.

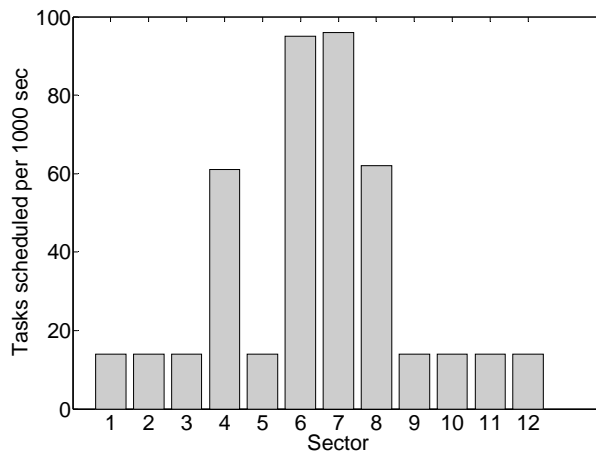


**Figure 76. Scheduling Algorithm for Scenarios with Targets Concentrated in Center (Dots Represent Targets)**

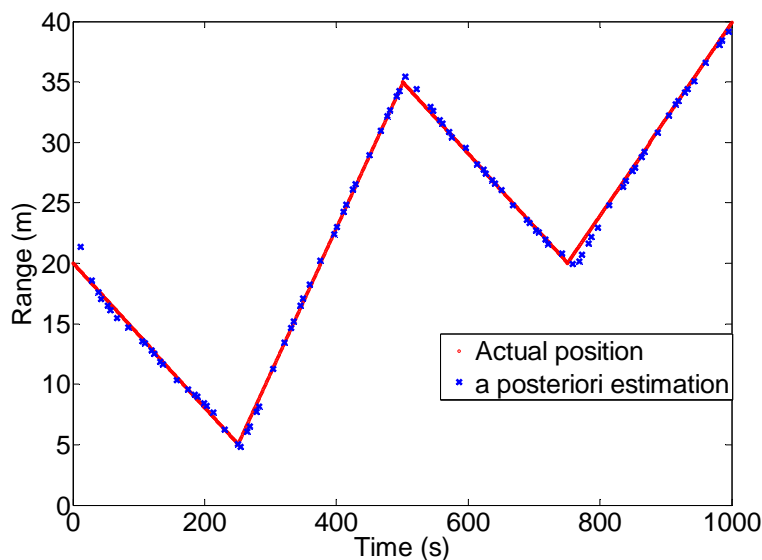
We used the particle filter algorithm to compare the tracking performance of our scheduling algorithms. 500 particles are used for the simulation results that follow. They are initially distributed uniformly over a range of 50 m. Periodically, the low weight particles are discarded and replaced with new particles which are distributed in range; then the weights are equalized.



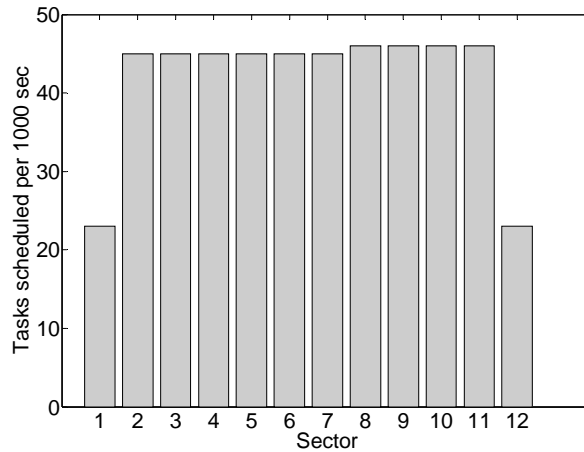
With three partial passes between each complete pass, the number of scheduled visits for all sectors and tracking results for sector 8 are shown in Figure 77 and Figure 78 respectively. We see that this algorithm gives us good number of visits for sectors being tracked. We also get very low tracking error even when there are abrupt changes in target direction as shown in Figure 78. We compare the results with a uniform algorithm which only makes complete passes and scans every sector in each direction. The results for this algorithm are shown in Figure 79 and Figure 80 respectively. Here we see that there is substantial error when the target motion has abrupt changes and it takes 3–4 visits for the tracker to correct itself.



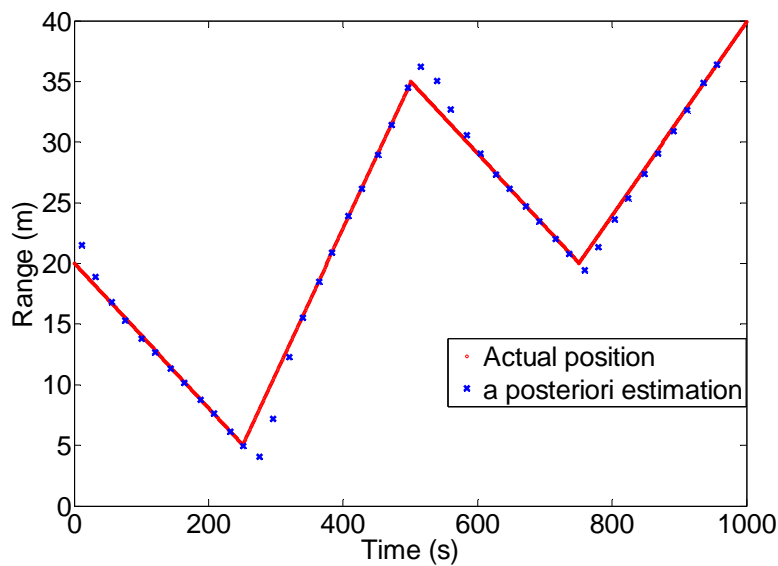
**Figure 77. Scheduling Results with 3 Partial Passes Between Complete Passes**



**Figure 78. Tracking Results for Sector 8 with 3 Partial Passes Between Complete Passes.**



**Figure 79. Scheduling Results with Complete Passes Only**

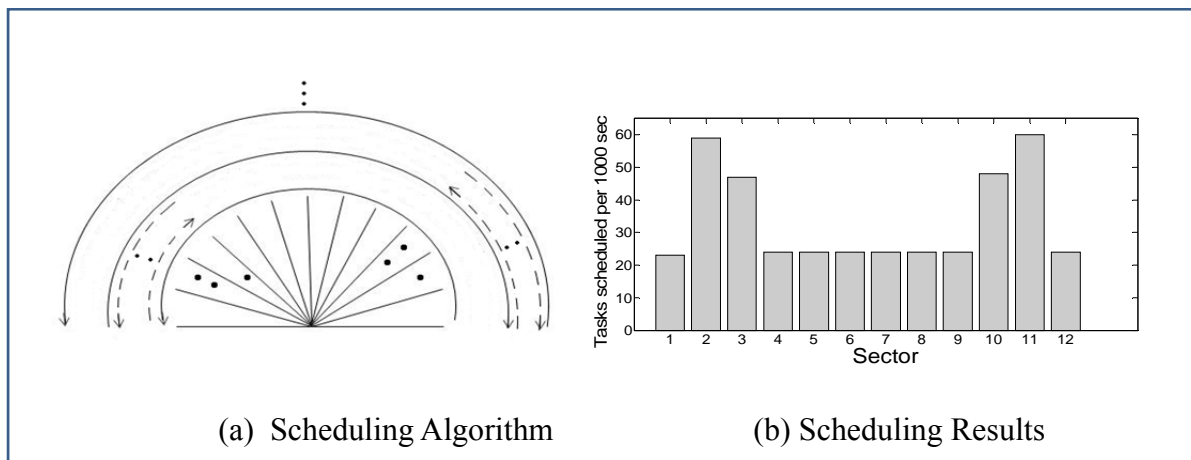


**Figure 80. Tracking Results with Complete Passes Only (Every Sector)**

#### 6.3.6.2. Targets on the Edges

For such target scenarios, the algorithm shown in Figure 81(a) is used. Partial passes are employed close to the edges to increase the number of tasks scheduled for sectors with targets. Only the sectors with targets are sampled in passes shown with dotted arcs. All sectors are sampled in passes shown with full arcs. After one complete cycle, the number of visits is measured for the sectors with targets and the number of partial passes between full passes is increased if required. For targets in sectors 2, 3, 10, and 11, the number of visits for 1000

seconds is as shown in Figure 81(b) when there are two partial passes between each complete pass.



**Figure 81. Scheduling for Scenarios with Targets Concentrated at the Edges**

From Figure 81(b), we observe that we achieve higher update rates for the sectors containing the targets.

## 6.4. Field Measurement Results

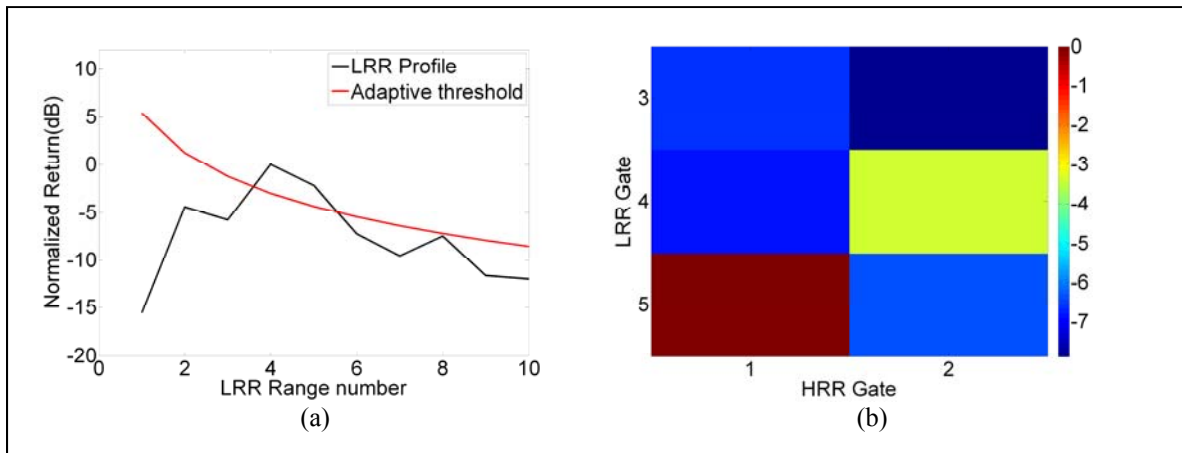
### 6.4.1. Field Measurement Results in Staring Mode

#### Target Scenario 6

Table 17 shows a target scenario 6 with 1 target. The LRR pass identifies LRR Gates 4 and 5 for further processing as shown in Figure 82(a). When we take HRR profiles for a particular LRR gate, we also include its neighboring gates to get a good idea of the target scene. Then Pass 2 is executed and now we get 3-dB separation between the peak and its neighboring gates as shown in Figure 82(b). Hence, no further passes are required.

**Table 17. Target Scenario 6**

Target Number	Range (m)	LRR Gate	Relative RCS
1	15.2	5	1



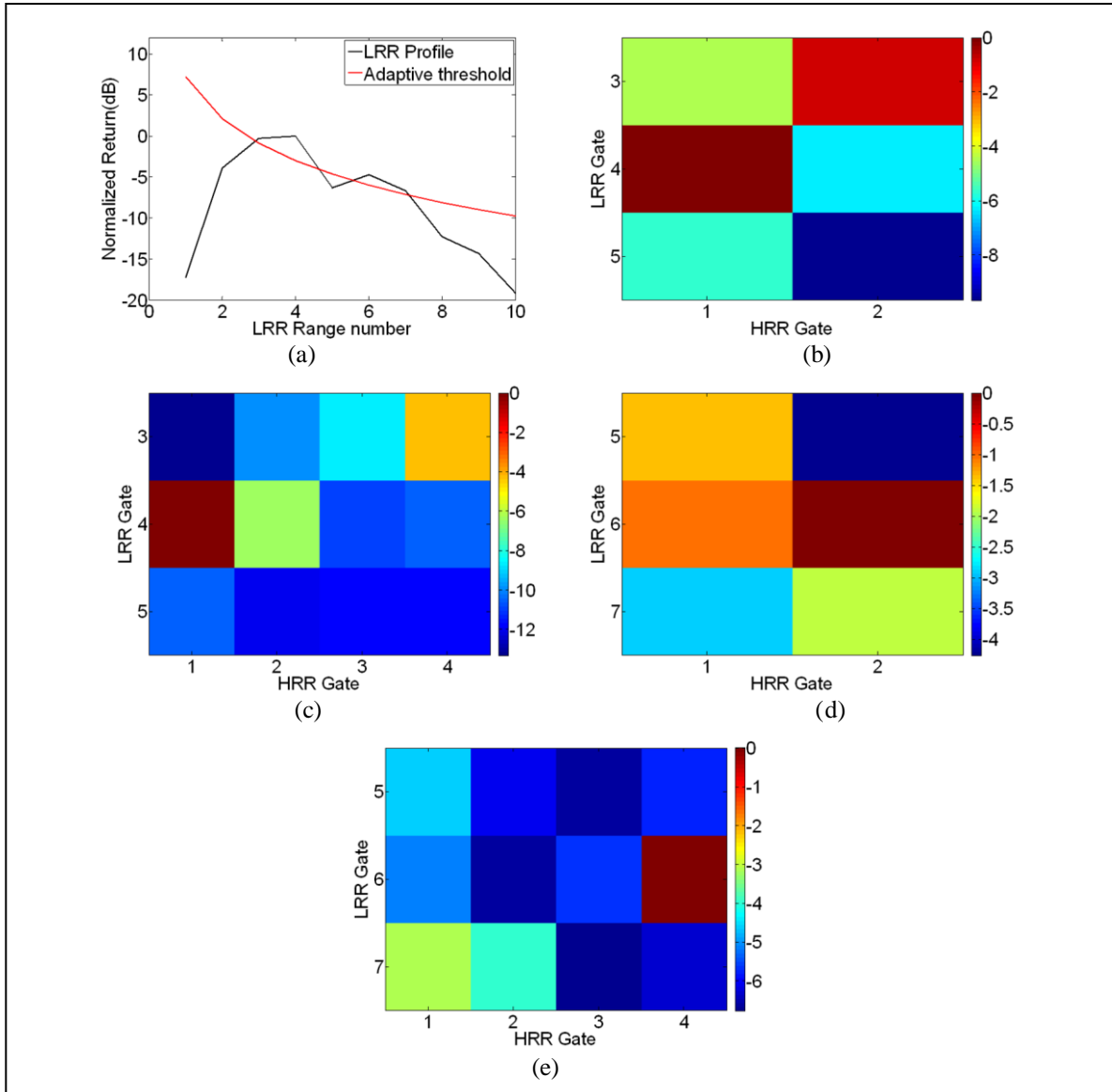
**Figure 82. Imaging Results for Target Scenario 6**  
**(a) Pass 1. (b) Pass 2, LRR Gate 4**

### Target Scenario 7

Table 18 shows a target scenario 7 with 2 targets. The LRR pass identifies LRR Gates 4 and 6 for further processing. For LRR gate 4, 3 dB separation is not obtained for LRR gate 4 as seen in Figure 83(b). Hence algorithm proceeds to Pass 3, where the separation is achieved and the algorithm stops. This is shown in Figure 83(c). Similarly the algorithm also stops at Pass 3 for LRR gate 6 as seen in Figure 83(e). These two targets are far apart and it may be argued that they should have been resolved adequately by the 80 MHz waveform itself. However, if a target is present on the boundary of a range gate, its response will not be contained in a single gate, leading to another pass.

**Table 18. Target Scenario 7**

Target Number	Range (m)	LRR Gate	Relative RCS
1	11.5	4	1
2	21.1	6	3



**Figure 83. Imaging Results for Target Scenario 7**

**(a) Pass 1. (b) Pass 2, LRR Gate 4. (c) Pass 3, LRR Gate 4. (d) Pass 2, LRR Gate 6. (e) Pass 3, LRR Gate 6**

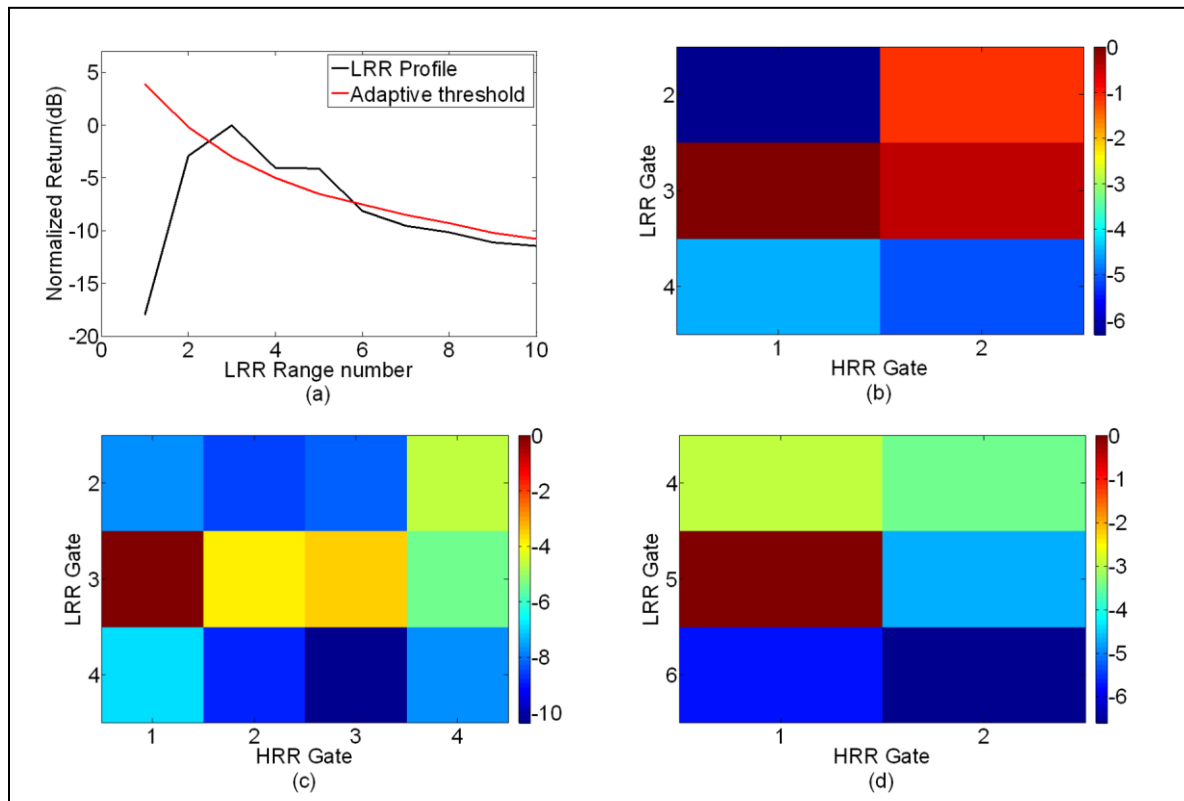
## Target Scenario 8

Table 19 shows a target scenario 8 with 3 targets. The LRR pass correctly identifies LRR Gates 3 and 5 for further processing. The multimodal algorithm runs and stops at Pass 3 for LRR Gate 3 and Pass 2 for LRR Gate 5, as shown in Figure 84. LRR Gate 3 contains 2 targets separated by 2.5 m. A resolution of about 1.25 m would be ideal. The multimodal radar required a bandwidth of 160 MHz for this target separation which corresponds to a resolution of 0.93 m. This is close to the expected result. Since 80 MHz is not sufficient to resolve these targets, we do not see

sufficient separation between cells in Pass 2 in Figure 84(b). Pass 3 gives us 3-dB separation as seen in Figure 84(c) and the algorithm stops. LRR Gate 5 contained a single target which resulted in the multimodal radar stopping at the first higher resolution pass of 80 MHz as shown in Figure 84(d).

**Table 19. Target Scenario 8**

Target Number	Range (m)	LRR Gate	Relative RCS
1	8.2	3	1
2	10.7	3	1
3	15.8	5	1



**Figure 84. Imaging Results for Target Scenario 8**

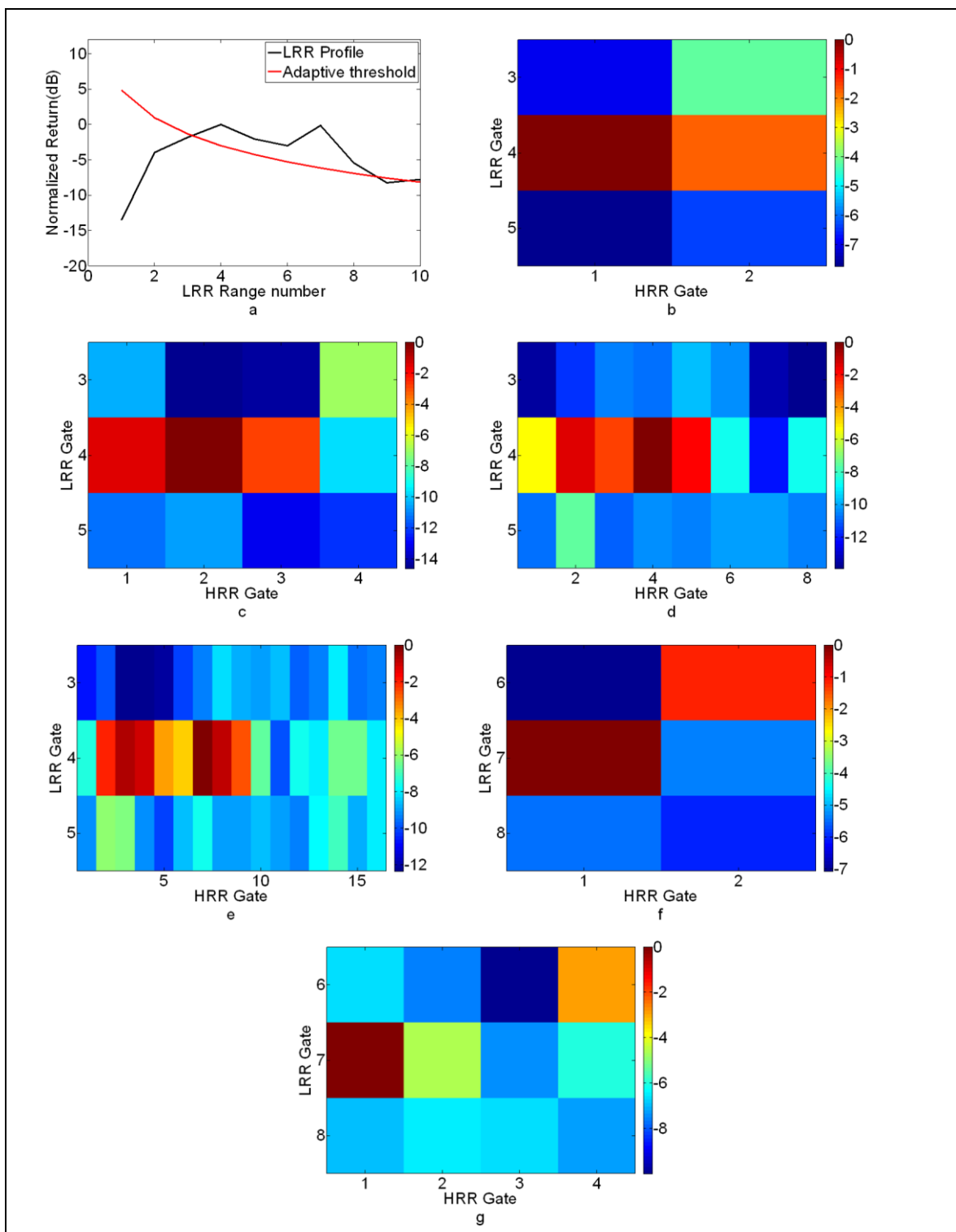
(a) Pass 1. (b) Pass 2, LRR Gate 3. (c) Pass 3, LRR Gate 3. (d) Pass 2, LRR Gate 5

## Target Scenario 9

Table 20 shows a target scenario 9 with three targets. A relative RCS of 4 (target number 3) indicates that four corner reflectors were placed side-by-side in that particular range cell. Here the algorithm stops at Pass 5 for LRR Gate 4 and Pass 3 for LRR Gate 7, as shown in Figure 85. The targets in LRR Gate 4 are separated by 0.6 m. Hence, in this case, a resolution of about 0.3 m would be ideal. The multimodal algorithm uses the entire bandwidth of 640 MHz corresponding to 0.23-m resolution for this target scene. This is close to the expected result. As seen in Figure 85(b) to Figure 85(d), none of the HRR passes from 2 to 4 is able to resolve these targets. HRR pass 5 clearly discerns the two targets in Figure 85(e). For LRR Gate 7, we had multiple corner reflectors placed at the same range to act as a single target. The radar required a bandwidth of 160 MHz for resolving this target as shown in Figure 85(d). While this is more than the expected bandwidth of 80 MHz for a single target, it may be due to the fact that our corner reflectors are not point targets and may happen to be split across two gates.

**Table 20. Target Scenario 9**

Target Number	Range (m)	LRR Gate	Relative RCS
1	11.5	4	1
2	12.1	4	1
3	22.1	7	4



**Figure 85. Imaging Results for Target Scenario 9**

**(a) Pass 1. (b) Pass 2, LRR Gate 4. (c) Pass 3, LRR Gate 4. (d) Pass 4, LRR Gate 4. (e) Pass 5, LRR Gate 4. (f) Pass 2, LRR Gate 7. (g) Pass 3, LRR Gate 7**

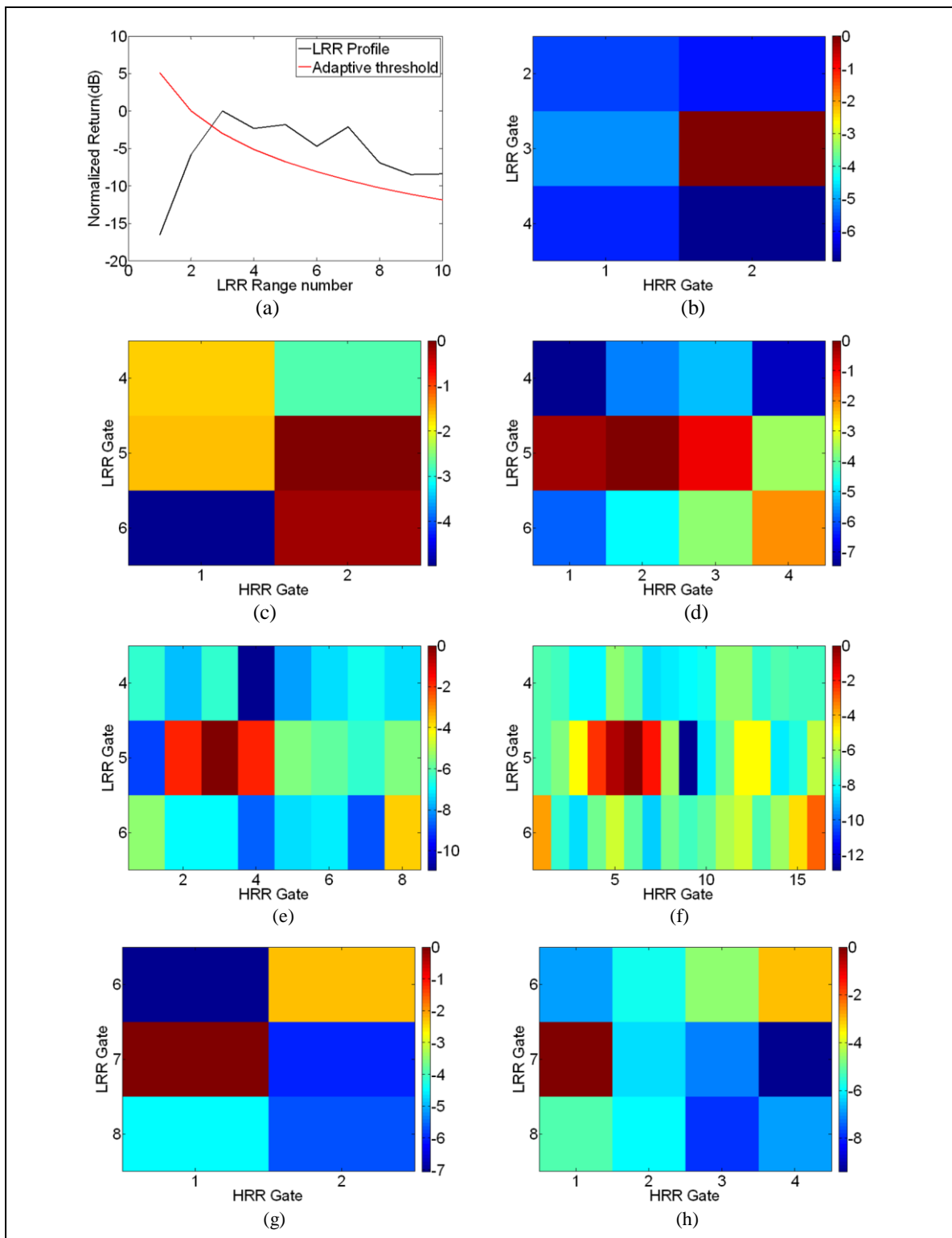


## Target Scenario 10

Table 21 shows a target scenario 10 with four targets. A relative RCS of 4 (target number 3) indicates that four corner reflectors were placed side-by-side in that particular range cell. Here the algorithm stops at Pass 2 for LRR Gate 3, Pass 5 for LRR Gate 5 and Pass 3 for LRR Gate 7 as shown in Figure 86. The targets in LRR Gate 5 are separated by 0.6 m. Hence, in this case, a resolution of about 0.3 m would be ideal. The multimodal algorithm uses the entire bandwidth of 640 MHz corresponding to 0.23-m resolution for this target scene. This is close to the expected result. As seen in Figure 86(c) to Figure 86(e), none of the HRR passes from 2 to 4 is able to resolve these targets. HRR pass 5 clearly discerns the two targets in Figure 86(f). For the targets in LRR Gates 3 and 7, the algorithm stops at pass 2 and pass 3 respectively.

**Table 21. Target Scenario 10**

<b>Target Number</b>	<b>Range (m)</b>	<b>LRR Gate</b>	<b>Relative RCS</b>
1	9.7	3	1
2	15.2	5	1
3	15.8	5	1
4	22.4	7	4



**Figure 86. Imaging Results for Target Scenario 10**

**(a) Pass 1. (b) Pass 2, LRR Gate 3. (c) Pass 2, LRR Gate 5. (d) Pass 3, LRR Gate 5. (e) Pass 4, LRR Gate 5. (f) Pass 5, LRR Gate 5. (g) Pass 2, LRR Gate 7. (h) Pass 3, LRR Gate 7**

### 6.4.2. Field Measurement Results in Scanning Mode

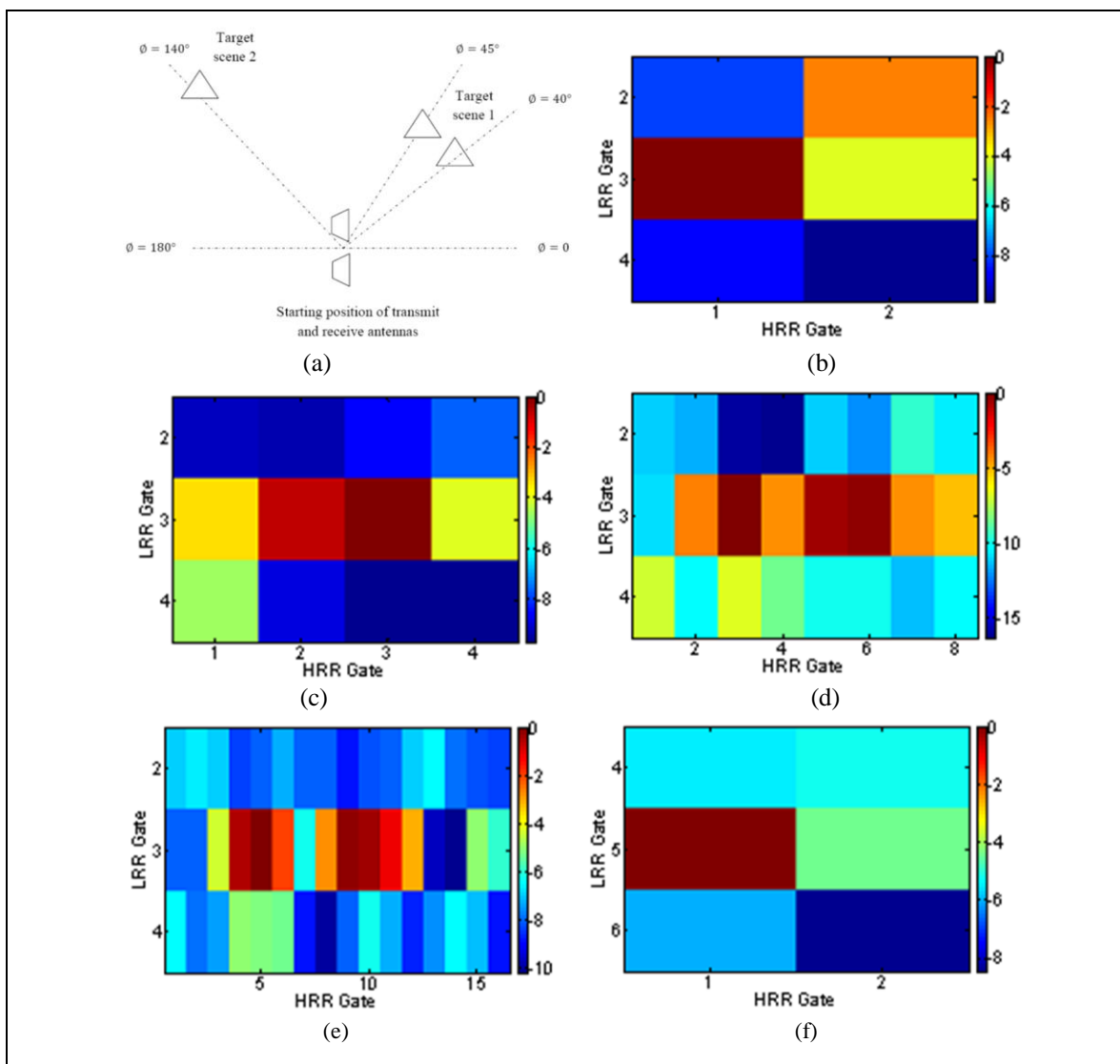
Measurements were also taken for scenarios where multiple target scenes are laid out in azimuthal fashion.

#### Target Scenario 11

Table 22 lists a target scenario 11 consisting of 2 target scenes. It is shown diagrammatically in Figure 87(a). LRR profiles are taken while the antennas are rotated with a step-size of  $15^\circ$ . The targets are correctly picked out from the LRR tracks and the multimodal algorithm is applied to these. The HRR images for these target scenes are shown in Figure 87. For target scene 1, LRR Gate 3 contains 2 targets separated by 1 m. HRR passes 2 and 3 are not able to resolve these targets, as shown in Figure 87(b) and 33(c). However, the two targets are discerned in HRR passes 4 and 5 in Figure 87(d) and 33(e). For target scene 2, HRR pass 2 is able to resolve the single target in LRR gate 5, as shown in Figure 87(f). This illustrates how the multimodal radar can be used to scan an area for targets and use only as much bandwidth as absolutely required.

**Table 22. Target Scenario 11**

Target Scene Number	$\phi$ (degrees)	Range (m)	LRR Gate	Relative RCS
1	40	8.5	3	2
	45	9.5	3	2
2	140	15.8	5	4



**Figure 87. HRR Images for Target Scenario 11**

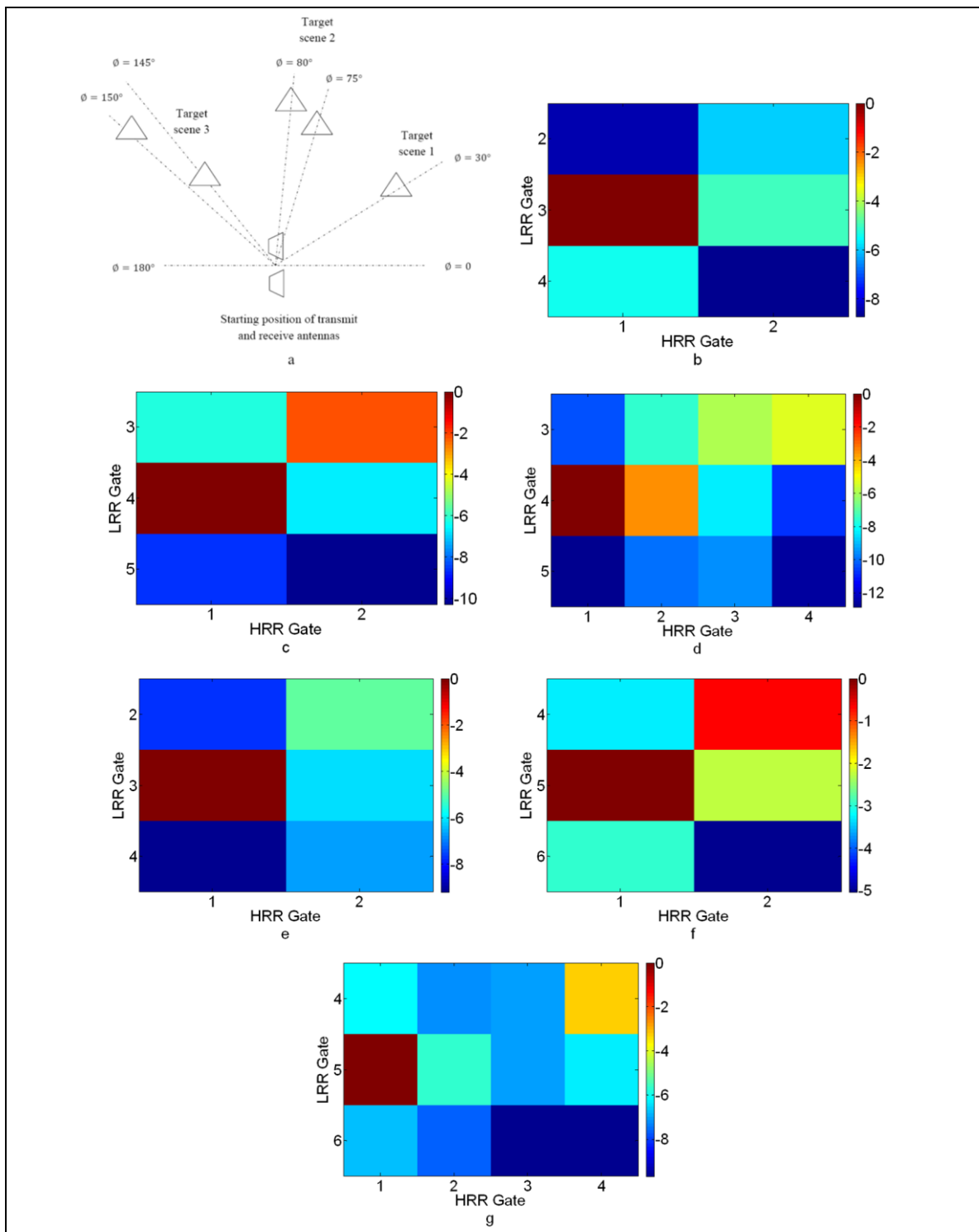
**(a) Diagrammatic representation. (b) Pass 2, Target scene 1. (c) Pass 3, Target scene 1. (d) Pass 4, Target scene 1. (e) Pass 5, Target scene 1. (f) Pass 2, Target scene 2**

## Target Scenario 12

Table 23 describes a target scenario 12 consisting of 3 target scenes. It is shown diagrammatically in Figure 88(a). Target scene 1 consists of a single target and is resolved by pass 2 as shown in Figure 88(b). Target scene 2 consists of 2 targets separated by 0.9 m. HRR pass 2 is not able to resolve these targets as seen in Figure 88(c). These 2 targets are discerned in Pass 3 as shown in Figure 88(d). Target scene 3 has one target each in LRR Gate 3 and 5. The target in LRR Gate 3 is resolved in Pass 2 as shown in Figure 88(e). The target in LRR Gate 5 in resolved in Pass 3 as shown in Figure 88(g). This again illustrates how the multimodal radar can be used for surveillance using low bandwidth waveform, switching to higher bandwidth waveforms when targets are detected.

**Table 23. Target Scenario 12**

Target Scene Number	$\phi$ (degrees)	Range (m)	LRR Gate	Relative RCS
1	30	8.2	3	1
2	75	11.3	4	1
	80	12.2	4	1
3	145	8.2	3	1
	150	15.2	5	2



**Figure 88. HRR Images for Target Scenario 12**

**(a) Diagrammatic representation. (b) Pass 2, Target scene 1. (c) Pass 2, Target scene 2. (d) Pass 3, Target scene 2. (e) Pass 2, Target scene 3, LRR Gate 3. (f) Pass 2, Target scene 3, LRR Gate 5. (g) Pass 3, Target scene 3, LRR Gate 5**

## 6.5. Conclusions

The methodology of a multimodal radar system with progressive resolution enhancement is described. This radar makes it possible to look at different target scenes with the appropriate bandwidth required to resolve the target features. It starts off with a view with lower resolution and stops when the appropriate resolution required for the target scene is reached. The saved bandwidth can be made available for use by other applications. It can also operate in scanning mode to provide surveillance capability over a wide area using a low bandwidth waveform and switching to other waveforms when required. The proof of concept was explained with scenarios for multimodal radar usage and simulation results.

The multimodal system radar system was built using standard components, arbitrary waveform generator and a digital oscilloscope. The waveform generator made it possible to use multiple waveforms. All the processing for multimodal algorithm was carried out using LabView and MATLAB. Surveillance function was enabled with the help of an antenna positioner and controller.

Experimental results were provided to give a demonstration of the multimodal radar algorithm in operation. For staring mode, we saw that the multimodal radar uses variable bandwidth based on the target scene. The scanning mode showed higher bandwidth usage only when targets are encountered. Thus, the system uses a field of view appropriate for the target scene.

Simulation results were shown to provide further insight into the performance characteristics of the radar. The ROC curves showed the superior performance of multimodal radar in terms of its lower probability of false alarm at higher values of probability of detection. We also saw the impact of a LRR pass of higher bandwidth. We also saw that the number of required passes increases as the target scene becomes more intricate or the SNR decreases.

A theoretical method was discussed to optimize the bandwidth required by the multimodal radar. Example scenarios were taken into consideration and it was shown that the results are close to the ones observed in the field. This further underscores the idea that each individual target scene requires an optimum bandwidth and there is not much gained by expending any additional bandwidth.

Several other considerations also come to mind. For example, there may be an overriding need to reserve a significant portion of the available spectrum for other applications, such as essential communications. In such a case, there may be an upper limit to the bandwidth available for the radar. In addition, if a specific smaller portion of the spectrum needs to be reserved for alternate applications, the radar may need to search for available contiguous spectrum for its operation within the entire band while avoiding the reserved subband. These issues require additional study. With the advent of software-defined RF technology, future multimodal radar systems can be designed to be reconfigurable, and therefore highly flexible and adaptive [121].

The concept of bandwidth sharing between multimodal radar and communications was introduced. The radar priority was calculated based on the target scene using fuzzy logic. Multi-objective optimization techniques were used to optimally share the bandwidth between radar and

communications. With the available bandwidth being scarce and costly, this is can be a great advantage for a multifunction system. Multiple applications can share the bandwidth and improve the total productivity of the system. Further study may be done into optimal allocation of bandwidth across different applications. It is also possible that one may not want to use up the entire bandwidth. Some bandwidth can be kept in reserve to be used when needed. Scheduling algorithms were suggested for the multimodal radar to increase the number of tasks scheduled for tracking and surveillance for specific target scenarios. Simulation results were provided to demonstrate the performance of these algorithms.

As RF systems become smaller and lighter to fit within smaller platforms, there is a need to integrate several functions into the same package for efficiency, such as surveillance, communications, tracking, and telemetry. In addition, bandwidth is a scarce commodity and must be managed wisely for optimal performance. Our research addresses the bandwidth management problem which leads to optimal scheduling of multifunctional RF systems.

Unambiguous target recognition is accomplished when all of the unique target features can be detected and matched against known templates. Separating the unique target scattering centers requires adequate bandwidth, which is inversely proportional to the separation between the scattering centers. Closely spaced target features require higher bandwidth compared to a target whose unique scattering centers are spaced far apart. In most cases, the highest bandwidth is employed, without assessing whether lower bandwidths may suffice for the specific target being surveilled. This wastes precious bandwidth which could be used for other applications. We have developed a methodology to start with a lower bandwidth and progressively increase the bandwidth until the target is clearly identified, at which point we do not increase the bandwidth further. Our approach is applicable to both staring and scanning mode scenarios and has been tested via simulation and experiments over a wide range of target scenarios. Freeing up unnecessary bandwidth can permit the multifunctional RF system to take advantage of the additional bandwidth for other important functions.

An important related issue is the allocation of bandwidth in multifunctional RF systems based upon priorities assigned by the commander. In most cases, a total bandwidth is assigned to the mission, and depending upon the priorities for each task, the bandwidth allocation may need to change dynamically for optimal performance. We have implemented a robust approach that employs fuzzy logic to perform multi-objective optimization and allocate bandwidth for different competing tasks, such as radar and communication.

In addition, we have also developed a strategy to optimize the tracking of targets by appropriate scheduling. For example, as the radar antenna scans in azimuth and identifies targets in angular sectors, the optimum solution is to perform tracking only over the sectors containing the targets. An approach has been developed to optimize the tracking instead of continuously tracking over all sectors, irrespective of target presence or absence. This technique will help higher frequency of target illumination and more revisits to ensure robust tracking.



## 7. Conclusions

The Integrated Multi-Modal RF Sensing project under the Multi-Yield Radio Frequency Countermeasures Investigations and Development contract for the AFRL Sensors Directorate performed basic research to examine key parameters of integrated multi-modal RF sensor design and algorithm development. Four Principal Investigators representing four different universities executed the program, with each PI being responsible for distinct but interrelated research areas. Dr. Muralidhar Rangaswamy of the Radio Frequency Exploitation Branch of the AFRL Sensors Directorate provided technical direction and collaboration throughout the project.

Dr. Steven Kay from the University of Rhode Island was responsible for researching the ***Statistical Framework*** for the overall IMMRF project. His research team investigated the problem of sensor integration to combine all the available information in a multi-sensor setting from a statistical standpoint. They studied the cases when the probability density function (PDF) is not completely known and proposed a novel method of constructing the joint PDF of the measurements from all the sensors based on the exponentially embedded family (EEF). The constructed PDF only requires knowledge of the joint PDF under a reference hypothesis and hence is useful in many practical cases. Examples and simulation results showed that the method requires less information compared to existing methods but attains comparable detection and classification performance. Additionally, the EEF can be used for robust signal detection.

Dr. Lee Swindlehurst from the University of California, Irvine led research into ***Algorithm Development and Implementation*** for the IMMRF project. His team addressed the problem of using multiple, spatially distributed, adaptive multi-modal sensors for multiple target tracking and data association. When left unmodeled or ignored, inherent field-of-view (FOV) limitations for each mode and sensor present a challenge to accurate tracking, appropriate sensor management, and system robustness. They proposed a relatively simple variance-penalty oriented modeling solution that effectively presents the FOV as a new design parameter. A novel closed-loop adaptive mode-parameter selection algorithm incorporating this penalty model was studied for use in both ideal and cluttered sensing environments. Simulations demonstrate the necessity and benefits of the proposed model and algorithm in a variety of tracking and association contexts as compared to completely unmodeled/non-adaptive methods.

Dr. Antonia Papandreou-Suppappola from the Arizona State University investigated ***Tracking and Waveform Agility***. Her team investigated the asymmetrical multi-modal sensing system for tracking low observable targets using RF and EO sensors, and developed the particle filter based recursive track before detect (TBD) algorithm for joint RF-EO multi-modal tracking to avoid loss of information caused by matched filter thresholding at low SNR. The TBD uses unthresholded matched filter measurements to estimate the target state, and performs tracking while detecting. They integrated the TBD approach with waveform-agile sensing as the TBD can directly incorporate the transmit waveform in its formulation. This allows for an adaptive selection of the transmit waveform at each time step without any restrictions on the SNR level. Her team also integrated the embedded exponential family (EEF) approach into the TBD algorithm in order to approximate joint distributions of dependent RF-EO measurements to further improve tracking performance, and used the IMMRF radar Test-Bed to evaluate the proposed algorithms. They also extended the TBD tracker to low observable multiple targets.

The new algorithm estimates the joint posterior probability density of all the target trajectories while keeping track of targets entering and leaving the noisy radar scene under observation using multiple modes. Multiple target tracking was investigated using the probability hypothesis density filter (PHDF) and integrated with adaptive waveform design.

Dr. Ram Narayanan from the Pennsylvania State University led the ***Radar Test-Bed Development*** efforts. His research team first studied the underlying methodology of an adaptive multimodal radar sensor that is capable of progressively varying its range resolution depending upon target scattering features. They developed a multimodal radar test-bed that enables the generation of linear frequency modulated waveforms of various bandwidths for achieving the optimum resolution to image the target. A theoretical method to optimize the bandwidth used by the radar was also developed. Novel bandwidth sharing scenarios between radar and communications functionalities were developed in which the surveillance space is divided into sectors and fuzzy logic is used to arrive at priorities for each sector. Priority for radar is based on separation between targets, signal to noise ratio, and existence of clutter. Finally, they considered the problem of optimal scheduling between tracking and surveillance for the multimodal radar, and developed algorithms for maximizing the track update rate and surveillance rate.

## 8. References

- [1] S.C.A. Thomopoulos, R. Viswanathan, and D.K. Bougoulas, "Optimal distributed decision fusion," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 25, pp. 761–765, Sep. 1989.
- [2] Z. Chair and P.K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 22, pp. 98–101, Jan. 1986.
- [3] A. Sundaresan, P.K. Varshney, and N.S.V. Rao, "Distributed detection of a nuclear radioactive source using fusion of correlated decisions," in *Information Fusion, 2007 10th International Conference on*, 2007, pp. 1–7.
- [4] S.G. Iyengar, P.K. Varshney, and T. Damarla, "A parametric copula based framework for multimodal signal processing," in *ICASSP*, 2009, pp. 1893–1896.
- [5] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley and Sons, second edition, 2006.
- [6] S. Kay, "Exponentially embedded families - new approaches to model order estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, pp. 333–345, Jan. 2005.
- [7] S. Kullback, *Information Theory and Statistics*, Courier Dover Publications, second edition, 1997.
- [8] S. Kay, A. Nuttall, and P.M. Baggenstoss, "Multidimensional probability density function approximations for detection, classification, and model order selection," *IEEE Trans. Signal Process.*, vol. 49, pp. 2240–2252, Oct. 2001.
- [9] L.D. Brown, *Fundamentals of Statistical Exponential Families*, Institute of Mathematical Statistics, 1986.
- [10] P.J. Bickel and K.A. Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics*, vol. 1, Pearson Prentice Hall, 2006.
- [11] E.L. Lehmann and J.P. Romano, *Testing Statistical Hypotheses*, Springer, third edition, 2005.
- [12] J.J. Higgins, "Some surface integral techniques in statistics," *The American Statistician*, vol. 29, pp. 43–46, Feb. 1975.
- [13] J. Pfanzagl and W. Wefelmeyer, *Contributions to a General Asymptotic Statistical Theory*, vol. 13 of *Lecture Notes in Statistics*, Springer-Verlag, 1982.
- [14] S. Kay, "Asymptotically optimal approximation of multidimensional pdf's by lower dimensional pdf's," *IEEE Trans. Signal Process.*, vol. 55, pp. 725–729, Feb. 2007.
- [15] D.G. Luenberger, *Linear and Nonlinear Programming*, Springer, second edition, 2003.
- [16] S. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1998.
- [17] S. Thomopoulos, R. Viswanathan, and D. Bougoulas, "Optimal distributed decision fusion," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 25, pp. 761–765, Sep. 1989.

- [18] Z. Chair and P. Varshney, "Optimal data fusion in multiple sensor detection systems," IEEE Trans. Aerosp. Electron. Syst., vol. 22, pp. 98–101, Jan. 1986.
- [19] J. Kittler, M. Hatef, R. Duin, and J. Matas, "On combining classifiers," IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, pp. 226–239, Mar. 1998.
- [20] A. Sundaresan, P. Varshney, and N. Rao, "Distributed detection of a nuclear radioactive source using fusion of correlated decisions," in Information Fusion, 2007 10th International Conference on, 2007, pp. 1–7.
- [21] S. Iyengar, P. Varshney, and T. Damarla, "A parametric copula based framework for multimodal signal processing," in ICASSP, 2009, pp. 1893–1896.
- [22] S. Kay and Q. Ding, "Exponentially embedded families for multimodal sensor processing," in ICASSP, Mar. 2010, pp. 3770–3773.
- [23] S. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [24] S. Kay, A. Nuttall, and P. Baggenstoss, "Multidimensional probability density function approximations for detection, classification, and model order selection," IEEE Trans. Signal Process., vol. 49, pp. 2240–2252, Oct. 2001.
- [25] T. Cover and J. Thomas, Elements of Information Theory, 2nd ed. John Wiley and Sons, 2006.
- [26] M. Westover, "Asymptotic geometry of multiple hypothesis testing," IEEE Trans. Inf. Theory, vol. 54, no. 7, pp. 3327–3329, Jul. 2008.
- [27] S. Kullback, Information Theory and Statistics, 2nd ed. Courier Dover Publications, 1997.
- [28] B. Efron, "Defining the curvature of a statistical problem (with applications to second order efficiency)," The Annals of Statistics, vol. 3, no. 6, pp. 1189–1242, 1975.
- [29] L. Brown, Fundamentals of Statistical Exponential Families. Institute of Mathematical Statistics, 1986.
- [30] S. Boyd and L. Vandenberghe, Convex Optimization. Cambridge University Press, 2004.
- [31] D. Luenberger, Linear and Nonlinear Programming, 2nd ed. Springer, 2003.
- [32] P. Hall, The bootstrap and Edgeworth expansion. Springer, 1997.
- [33] S. Kay, Q. Ding, and D. Emge, "Joint pdf construction for sensor fusion and distributed detection," in International Conference on Information Fusion, Jun. 2010.
- [34] S. A. Kassam, Signal Detection in Non-Gaussian Noise. Springer-Verlag, 1988.
- [35] S. Kay, Q. Ding, and M. Rangaswamy, "Sensor integration for classification," in Asilomar Conference on Signals, Systems, and Computers, Nov. 2010.
- [36] S. Kay, "Exponentially embedded families - new approaches to model order estimation," IEEE Trans. Aerosp. Electron. Syst., vol. 41, pp. 333–345, Jan. 2005.
- [37] S. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1998.
- [38] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [39] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*, 3rd ed. Springer, 2005.

- [40] A. Sundaresan, P. Varshney, and N. Rao, "Distributed detection of a nuclear radioactive source using fusion of correlated decisions," in *Information Fusion, 2007 10th International Conference on*, 2007, pp.1–7.
- [41] S. Iyengar, P. Varshney, and T. Damarla, "A parametric copula based framework for multimodal signal processing," in *ICASSP*, 2009, pp. 1893–1896.
- [42] S. Kay and Q. Ding, "Exponentially embedded families for multimodal sensor processing," in *ICASSP*, Mar. 2010.
- [43] S. Kay, Q. Ding, and D. Emge, "Joint PDF construction for sensor fusion and distributed detection," in *International Conference on Information Fusion*, Jun. 2010.
- [44] L. Brown, *Fundamentals of Statistical Exponential Families*. Institute of Mathematical Statistics, 1986.
- [45] S. Kullback, *Information Theory and Statistics*, 2nd ed. Courier Dover Publications, 1997.
- [46] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley and Sons, 2006.
- [47] J. Pfanzagl and W. Wefelmeyer, *Contributions to a General Asymptotic Statistical Theory*, ser. *Lecture Notes in Statistics*. Springer-Verlag, 1982, vol. 13.
- [48] P. Bickel and K. Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics*. Pearson Prentice Hall, 2006, vol. 1.
- [49] G. Grimmett and D. Stirzaker, *Probability and Random Processes*. Oxford University Press, 2001.
- [50] J. Kittler, M. Hatef, R. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, pp. 226–239, Mar. 1998.
- [51] E. Blasch and C. Yang, "Ten methods to fuse GMTI and HRRR measurements for joint tracking and identification," in *Proceedings of the 7th International Conference on Information Fusion (FUSION'04)*, 2004, pp. 1006-1013.
- [52] D. J. Kershaw and R. J. Evans, "Optimal waveform selection for tracking systems," *IEEE Transactions on Information Theory*, vol. 40, no. 5, pp. 1536-1550, 1994.
- [53] D. J. Kershaw and R. J. Evans, "Waveform selective probabilistic data association," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 4, pp. 1180-1188, Oct. 1997.
- [54] S. P. Sira, Y. Li, A. Papandreou-Suppappola, D. Morrell, D. Cochran, and M. Rangaswamy, "Waveform-agile sensing for tracking," *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 53-64, Jan. 2009.
- [55] J. J. Zhang, A. Papandreou-Suppappola, and M. Rangaswamy, "Multi-target tracking using multi-modal sensing with waveform configuration," in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2010, pp. 3890-3893.
- [56] S. A. Miller, Z. A. Harris, and E. K. P. Chong, "A POMDP Framework for Coordinated Guidance of Autonomous UAVs for Multitarget Tracking," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, pp. 1-18, Jan. 2009.
- [57] J. A. Fuemmeler and V. V. Veeravalli, "Smart Sleeping Policies for Energy Efficient Tracking in Sensor Networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 2091-2101, May 2008.

- [58] A. Kansal, W. Kaiser, G. Pottie, M. Srivastava, and G. Sukhatme, "Reconfiguration methods for mobile sensor networks," *ACM Transactions on Sensor Networks*, vol. 3, no. 4, p. 22-es, Oct. 2007.
- [59] A. Mittal and L. S. Davis, "A General Method for Sensor Planning in Multi-Sensor Systems: Extension to Random Occlusion," *International Journal of Computer Vision*, vol. 76, no. 1, pp. 31-52, Jul. 2007.
- [60] J. Zhao, S.-C. S. Cheung, and T. Ngyuen, "Optimal Visual Sensor Network Configuration," in *Multi-Camera Networks: Principles and Applications*, H. K. Aghajan and A. Cavallaro, Eds. Oxford: Academic Press, 2009, pp. 139-162.
- [61] F. Daum, "A System Approach to Multiple Target Tracking," in *Multitarget-Multisensor Tracking: Applications and Advances, Volume II*, Y. Bar-Shalom, Ed. Norwood, MA: Artech House, 1992, pp. 149-181.
- [62] J. P. Stralka and W. G. Fedarko, "Pulse Doppler Radar," in *Radar Handbook*, 3rd ed., no. 1990, M. I. Skolnik, Ed. McGraw-Hill, 2008, pp. 4.1-4.54.
- [63] J. M. Lloyd, "Fundamentals of Electro-Optical Imaging Systems Analysis," in *The Infrared & Electro-Optical Systems Handbook v. 4*, M. C. Dudzik, Ed. Ann Arbor, Michigan and Bellingham, Washington: Infrared Information Analysis Center and SPIE Optical Engineering Press, 1993, pp. 1-54.
- [64] L. Y. Pao, "Centralized multisensor fusion algorithms for tracking applications," *Control Engineering Practice*, vol. 2, no. 5, pp. 875-887, Oct. 1994.
- [65] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Norwood, MA: Artech House, 1999.
- [66] C. W. Frei and L. Y. Pao, "Alternatives to Monte-Carlo simulation evaluations of two multisensor fusion algorithms," *Automatica*, vol. 34, no. 1, pp. 103-110, Jan. 1998.
- [67] T. Fortmann, Y. Bar-Shalom, M. Scheffe, and S. Gelfand, "Detection thresholds for tracking in clutter--A connection between estimation and signal processing," *IEEE Transactions on Automatic Control*, vol. 30, no. 3, pp. 221-229, Mar. 1985.
- [68] S. Boyd and L. Vanderberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [69] K. Oldham, J. Myland, and J. Spanier, *An Atlas of Functions*. New York, NY: Springer New York, 2009, pp. 1-6.
- [70] D. W. Casbeer and A. L. Swindlehurst, "A Non-Search Optimal Control Solution for a Team of MUAVS in a Reconnaissance Mission," *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, no. 1, p. IV-1169-IV-1172, 2006.
- [71] C. M. Kreucher, K. Kastella, and A. O. Hero, "Sensor management using an active sensing approach," *Signal Processing*, vol. 85, no. 3, pp. 607-624, Mar. 2005.
- [72] P. Zhan, D. W. Casbeer, and A. L. Swindlehurst, "Adaptive Mobile Sensor Positioning for Multi-Static Target Tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 1, pp. 120-132, Jan. 2010.
- [73] E. Blasch and B. Kahler, "Multiresolution EO/IR target tracking and identification," *International Conference on Information Fusion*, pp. 275-282, July 2005.
- [74] A. Theil, L. Kester, S. P. van den Broek, P. van Dorp, and R. van Sweeden, "FRESNEL program: Fusion of radar and electro-optical signals for surveillance on land," *SPIE Conference*

on Signal Processing, Sensor Fusion, and Target Recognition, vol. 4380, pp. 453-461, August 2001.

[75] D. J. Salmond and H. Birch, "A particle filter for track-before-detect," *Proceedings American Control Conference*, vol. 5, pp. 3755-3760, 2001.

[76] Y. Boers and J. N. Driessen, "Multitarget particle filter track before detect application," *IEEE Proceedings - Radar, Sonar and Navigation*, vol. 151, pp. 351-357, 2004.

[77] S. J. Davey, B. Cheung, and M. G. Rutten, "Track-before-detect for sensors with complex measurements," *International Conference on Information Fusion*, pp. 619-625, July 2009.

[78] S. P. Sira, A. Papandreou-Suppappola, and D. Morrell, "Dynamic configuration of time-varying waveforms for agile sensing and tracking in clutter," *IEEE Transactions on Signal Processing*, vol. 55, pp. 3207-3217, July 2007.

[79] S. P. Sira, Y. Li, A. Papandreou-Suppappola, D. Morrell, D. Cochran, M. Rangaswamy, "Waveform-agile sensing for tracking: A review perspective," *IEEE Magazine on Signal Processing*, vol. 26, pp. 53-64, 2009.

[80] J. J. Zhang, A. Papandreou-Suppappola, and M. Rangaswamy, "Multitarget tracking using multi-modal sensing with waveform configuration," *IEEE International Conference on Acoustics Speech and Signal Processing*, pp. 3890-3893, March 2010.

[81] J. J. Zhang, Q. Ding, S. Kay, A. Papandreou-Suppappola, M. Rangaswamy, "Agile multi-modal tracking with dependent measurements," *Asilomar Conference on Signals, Systems, and Computers*, pp. 1653-1657, November 2010.

[82] S. Liu, S. Bhat, J. J. Zhang, Q. Ding, S. Kay, A. Papandreou-Suppappola, M. Rangaswamy, "Design and performance of an integrated waveform agile multi-modal track-before-detect sensing system," *Asilomar Conference on Signals, Systems, and Computers*, pp. 1530-1534, November, 2011.

[83] S. Liu, "Integrated waveform-agile multi-modal track-before-detect algorithms for tracking low observable targets," Master's thesis, Arizona State University, May 2012.

[84] R. Piwowarski, B. O'Donnell, J. J. Zhang, A. Papandreou-Suppappola, "Waveform-Agile Track-Before-Detect for Low Observable Targets," *International Waveform Diversity and Design Conference*, January 2012.

[85] S. Kay, "Exponentially embedded families - new approaches to model order estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 4, pp. 333-345, January 2005.

[86] S. Kay and Q. Ding, "Exponentially embedded families for multimodal sensor processing," *IEEE International Conference on Acoustics Speech and Signal Processing*, pp. 3770-3773, March 2010.

[87] Y. Bar-Shalom and X.-R. Li, *Multisensor Tracking: Principles and Techniques*. New York: YBS Publishing, 1995.

[88] M. Taj and A. Cavallaro, "Multi-camera track-before-detect," *IEEE International Conference on Distributed Smart Cameras*, pp. 1-6, September 2009.

[89] S. Ebenezer and A. Papandreou-Suppappola, "Efficient multiple particle filtering for track-before-detect target tracking," *International Waveform Diversity and Design Conference*, January 2012.

- [90] Y. Bar-Shalom, F. Daum, and J. Huang, "The probabilistic data association filter," *IEEE Control Systems*, vol. 29, pp. 82-100, December 2009.
- [91] R. P. S. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, pp. 1152-1178, October 2003.
- [92] B. O'Donnell, J. J. Zhang, A. Papandreou-Suppappola, M. Rangaswamy, "Waveform-agile multiple target tracking using probability hypothesis density filtering," *IEEE Sensor Array and Multichannel Signal Processing Workshop*, June 2012.
- [93] S. Sarkka, A. Vehtari, and J. Lampinen, "Rao-Blackwellized particle filter for multiple target tracking," *Information Fusion*, vol. 8, no. 1, pp. 2-15, 2007.
- [94] C. Kreucher, K. Kastella, and A. O. Hero III, "Multitarget tracking using the joint multitarget probability density," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 4, pp. 1396-1414, October 2005.
- [95] B.-N. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo implementation of the PHD filter for multi-target tracking," *International Conference of Information Fusion*, vol. 2, pp. 792-799, 2003.
- [96] Narayanan, R.M., "Aspects of radar waveform design," *Proc. International Radar Symposium India (IRSI-09)*, Bangalore, India, December 2009, pp. 8-14.
- [97] Manners, D.M., "ART – An adaptive radar testbed," *Proc. IEE Colloquium on Real-Time Management of Adaptive Radar Systems*, London, UK, June 1990, pp. 6/1-6/7.
- [98] Adler, E., Clark, J., Conn, M., Phu, P., and Scheiner, B., "Low-cost technology for multimode radar," *IEEE AES Systems Magazine*, 14, (6), June 1999, pp. 23-27.
- [99] Huizing, A.G., "Wideband vs. multiband trade-offs for a scalable multifunction RF system," *Proc. 2005 IEEE International Radar Conference*, Arlington, VA, May 2005, pp. 155-160.
- [100] Hughes, P.K., and Choe, J.Y., "Overview of advanced multifunction RF system (AMRFS)," *Proc. IEEE Conference on Phased Array Systems and Technology*, Dana Point, CA, May 2000, pp. 21-24.
- [101] Huizing, A.G., Otten, M.P.G., van Rossum, W.L., van Dijk, R., Maas, A.P.M., van der Houwen, E.H., and Bolt, R.J., "Compact scalable multifunction RF payload for UAVs with FMCW radar and ESM functionality," *Proc. 2009 International Radar Conference, Bordeaux, France*, October 2009.
- [102] Kahrilas, P.J., "HAPDAR—An operational phased array radar," *Proc. of the IEEE*, 56, (11), November 1968, pp. 1967-1975.
- [103] Moore, A.R., Salter, D.M., and Stafford, W.K., "MESAR (Multi-function, electronically scanned, adaptive radar)," *Proc. Radar Systems*, Edinburgh, UK, October 1997, pp. 55-59.
- [104] Groger, I., Sander, W., and Wirth, W.D., "Experimental phased array radar ELRA with extended flexibility," *IEEE Aerospace and Electronic Systems Magazine*, 5, (11), November 1990, pp. 26-30.
- [105] Smits, A.B., and van Genderen, P., "The APAR multifunction radar – System overview," *Proc. IEEE International Symposium on Phased Array Systems and Technology*, Boston, MA, October 2003, pp. 241-246.



- [106] Capraro, G.T., Farina, A., Griffiths, H., and Wicks, M.C., "Knowledge-based radar signal and data processing: a tutorial review," *IEEE Signal Processing Magazine*, 23, (1), January 2006, pp. 18–29.
- [107] Miranda, S.L.C., Baker, C.J., Woodbridge, K., and Griffiths, H.D., "Comparison of scheduling algorithms for multifunction radar," *IET Radar Sonar and Navigation*, 1, (6), 2007, pp. 414–424.
- [108] Haykin, S., "Cognitive radar: A way of the future," *IEEE Signal Processing Magazine*, 23, (1), Jan. 2006, pp. 30–40.
- [109] Bell, M. R., "Information theory and radar waveform design," *IEEE Transactions on Information Theory*, 39, (5), Sep. 1993, pp. 1578–1597.
- [110] Jackson, C.A., Holloway, J.R., Pollard, R., Larson, R., Sarno, C., Woodbridge, K., Ormondroyd, R.F., Lewis, M.B., and Stove, A.G., "Spectrally efficient radar systems in the L and S bands," *Proc. 2007 IET International Conference on Radar Systems*, Edinburgh, UK, October 2007.
- [111] Wang, L., McGeehan, J., Williams, C., and Doufexi, A., "Application of cooperative sensing in radar-communications coexistence," *IET Communications*, 2, (6), June 2008, pp. 856–868.
- [112] Zhang, J.J., Papandreou-Suppappola, A., and Rangaswamy, M., "Multi-target tracking using multi-modal sensing with waveform configuration," *IEEE international conference on Acoustics Speech and Signal processing*, Dallas, Texas, March 2010, pp. 3890–3893.
- [113] Bhat, S.S., Narayanan, R.M., and Rangaswamy, M., "An adaptive multimodal radar system with progressive resolution enhancement," *Proc. 5<sup>th</sup> International Waveform Diversity and Design Conference*, Niagara Falls, Canada, August 2010, pp. 95–99.
- [114] Bhat, S.S., Narayanan, R.M., and Rangaswamy, M., "Design and performance of a multimodal radar test-bed for progressive resolution enhancement," *Proc. 45<sup>th</sup> Annual Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, California, Nov. 2010, pp. 1667–1671.
- [115] Boyd, S., and Vandenberghe, L., "Convex Optimization," (Cambridge Univ. Press, Cambridge, UK, 2004), pp. 215–216.
- [116] Mir, H.S., and Carlson, B.D., "On the definition of radar range resolution for targets of greatly differing RCS," *IEEE Transactions on Instrumentation and Measurement*, 61, (3), Mar. 2012, pp. 655–663.
- [117] Miranda, S.L.C., Baker, J., Woodbridge, K., and Griffiths, H.D., "Fuzzy logic approach for prioritization of radar tasks and sectors of surveillance in multifunction radar," *IET Radar Sonar and Navigation*, 1, (2), 2007, pp. 131–141.
- [118] Vincent, T.L., and Grantham, W.J., "Optimality in Parametric Systems," New York: John Wiley and Sons, 1981.
- [119] Puranik, S.P., and Tugnait, J.K., "On adaptive sampling for multisensor tracking of a maneuvering target using IMM/PDA filtering," in *Proc. 2005 American Control Conference*, Portland, OR, June 2005, pp. 1263–1268.
- [120] Lo, K.C., and Purvis, A., "Application of random and pseudo-random sampling for tracking moving objects," in *Proc. 1994 International Symposium on Speech, Image Processing and Neural Networks*, Hong Kong, 1994, pp. 260–263.

[121] Frankford, M.T., Majurec, N., and Johnson, J.T., “Software-defined radar for MIMO and adaptive waveform applications,” *Proc. 2010 IEEE Radar Conference*, Washington, DC, May 2010, pp. 724–728.

## **List of Acronyms**

1-D	One-Dimensional
2-D	Two-Dimensional
3-D	Three-Dimensional
AF	Ambiguity Function
AMRFS	Advanced Multifunction RF System
AWG	Arbitrary Waveform Generator
CBRN	Chemical, Biological, Radiological, And Nuclear
CDF	Cumulative Distribution Function
CRLB	Cramer-Rao Lower Bound
EEF	Exponentially Embedded Family
EEF	Embedded Exponential Family
EM	Electromagnetic
EO	Electro-Optical
FOV	Field Of View
GLRT	Generalized Likelihood Ratio Test
GMTI	Ground Moving Target Indicator
HRR	High-Range Resolution
IID	Independent Identically Distributed
IMMRF	Integrated Multi-Modal RF
JPDA	Joint Probabilistic Data Association
KL	Kullback-Leibler
LRR	Low-Range Resolution
MAP	Maximum Posteriori Probability
MESAR	Multi-Function Electronically Scanned, Adaptive Radar
MHT	Multiple Hypothesis Tracker
MLE	Maximum Likelihood Estimates
MSE	Mean-Squared Error
MTT	Multiple Target Tracking
MYRIAD	Multi-Yield RFCM Investigations and Development
NP	Neyman-Pearson Rule
PCRLB	Posterior Cramer-Rao Lower Bound
PD	Probability Of Detection
PDF	Probability Density Function
PF	Particle Filter
PFA	Probability Of False Alarm
PF-TBD	Particle Filter Track Before Detect
PHDF	Probability Hypothesis Density Filter
RCS	Radar Cross Section
RF	Radio Frequency

RFCM	Radio Frequency Counter Measures
ROC	Receiver Operating Characteristic
SMC	Sequential Monte Carlo
SNR	Signal-To-Noise Ratio
TBD	Track-Before-Detect
UMP	Uniformly Most Powerful